线性代数知识整理（浓缩版）

第一章 矩阵与线性方程组

一、向量及其基本性质

1、行向量与列向量的概念（略）

2、n维向量空间：$\boldsymbol{R}^n=\{\vec{a}\mid\vec{a}=(a_1,a_2,...,a_n)$，$a_i\in\boldsymbol{R}$，$i=1,2,...,n\}$

在这个空间上定义了向量的一系列运算：

（1）加法；（2）数乘；

（3）取模：$\|\boldsymbol{a}\|=\sqrt{\sum_{i=1}^{n}a_i^2}$；

（4）（在 n 维欧氏空间 $\mathbf{E}^n$ 上定义）内积（点乘，数量积）：$\boldsymbol{a}\cdot\boldsymbol{b}=\sum_{i=1}^{n}a_ib_i$.

二、矩阵及其运算

1、基本表示：$m\times n$矩阵$\boldsymbol{A}=(a_{ij})_{m\times n}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&&&\\a_{m1}&a_{n2}&\cdots&a_{mn}\end{pmatrix}$.

2、概念：实矩阵，复矩阵，方阵，主对角线，同型矩阵，零矩阵，单位矩阵，Kronecker 记号（δ_{ij}），对角矩阵，上（下）三角矩阵……

3、矩阵的运算

（1）加法：$\boldsymbol{A}+\boldsymbol{B}=(a_{ij}+b_{ij})_{m\times n}$（满足交换律、结合律）；

（2）数乘：$k\boldsymbol{A}=(ka_{ij})_{m\times n}$（满足分配律）；

（3）转置：$\boldsymbol{A}^{\mathrm{T}}=(a_{ji})_{n\times m}$（行变列，列变行），$(\boldsymbol{A}+\boldsymbol{B})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{B}^{\mathrm{T}}$，$(k\boldsymbol{A})^{\mathrm{T}}=k\boldsymbol{A}^{\mathrm{T}}$；

（4）共轭：$\overline{\boldsymbol{A}}=(\overline{a}_{ij})_{m\times n}$，共轭转置$\boldsymbol{A}^{\mathrm{H}}=(\overline{\boldsymbol{A}})^{\mathrm{T}}=\overline{(\boldsymbol{A}^{\mathrm{T}})}$；

（5）乘法

①条件：$\boldsymbol{AB}$ 存在的条件是 $\boldsymbol{A}$ 的列数等于 $\boldsymbol{B}$ 的行数。

②公式：对于$m\times n$矩阵$\boldsymbol{A}$与$n\times p$矩阵$\boldsymbol{B}$，$\boldsymbol{AB}=(\sum_{k=1}^{n}a_{ik}b_{kj})_{m\times p}$.$\boldsymbol{AB}$ 的第 i 行第 j 列的元素等于 $\boldsymbol{A}$ 的第 i 行与 $\boldsymbol{B}$ 的第 j 列作内积。（满足结合律和分配律，不满足交换律，左乘右乘不一样；$(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$）

4、分块矩阵及其运算（略，注意 $\boldsymbol{A}^{\mathrm{T}}=\begin{pmatrix}\boldsymbol{A}_{11}^{\mathrm{T}}&\boldsymbol{A}_{21}^{\mathrm{T}}\\\boldsymbol{A}_{12}^{\mathrm{T}}&\boldsymbol{A}_{22}^{\mathrm{T}}\end{pmatrix}$）

二、行列式

1、行列式的本质——与排列有关

（1）只有方阵才有行列式。

（2）排列的逆序数：设 $j_1,j_2,...,j_n$ 是 $1,2,...,n$ 的一个排列。将满足 $k<l$ 且 $j_k>j_l$ 的数对(k,l)的个数称为$j_1,j_2,...,j_n$ 的逆序数，记作$\tau(j_1,j_2,...,j_n)$。

（3）n 阶行列式的原始定义：

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_{j_1, j_2, \dots, j_n} (-1)^{\tau(j_1, j_2, \dots, j_n)} a_{1j_1} a_{2j_2} \dots a_{nj_n}.$$

（4）沿一行展开：

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_{j=1}^{n} a_{ij} A_{ij}，其中 A_{ij} = (-1)^{i+j} \Delta_{ij} 为代数余子式，\Delta_{ij} 为余子式。$$

Δ_{ij}即行列式中去掉第i行与第j列所得的行列式。（注意区分A_{ij}与Δ_{ij}）

沿一列展开：$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_{i=1}^{n} a_{ij} A_{ij}.$

*沿两行展开：$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_{1 \le j_1 < j_2 \le n} (-1)^{i_1+i_2+j_1+j_2} \begin{vmatrix} a_{i_1 j_1} & a_{i_1 j_2} \\ a_{i_2 j_1} & a_{i_2 j_2} \end{vmatrix} B_{i_1 i_2 j_1 j_2}$。

$B_{i_1 i_2 j_1 j_2}$是行列式中去掉第i_1、i_2行与第j_1、j_2列所得的行列式。（可类似地定义多行展开）

2、行列式的性质（$\boldsymbol{A}$ 为 n 阶方阵）

（1）$|\boldsymbol{A}| = |\boldsymbol{A}^{\mathrm{T}}|$；

（2）互换行列式的两行（列），行列式的值变号；

（3）若行列式有两行（列）成比例，则行列式的值为 0；

（4）两个<u>仅有一行（列）不同</u>的行列式的和，等于<u>将该行（列）相加</u>，其余行（列）不变的行列式；（对比：两个矩阵相加等价于把<u>所有对应元素</u>相加）

（5）★将行列式的某一行（列）乘以一个数加到另一行（列），行列式的值不变。

（6）将行列式乘以λ，等价于将其<u>某一行（列）</u>乘以λ。（对比：将矩阵乘以λ，等价于将其<u>所有元素</u>乘以λ；由此可得$|\lambda \boldsymbol{A}| = \lambda^n |\boldsymbol{A}|.$）

（7）三角形行列式的值等于主对角元素的乘积。

注：①在计算阶数已知的行列式时，先看是否有成比例的行（列），如果没有，则不断用其他性质进行化简，如果实在化不成对角行列式，应适时放弃变换，在某一行（列）0 元素较多时即可直接展开；

②对于阶数 n 未知的行列式，往往有一定规律（否则难以计算），一般来说需要所有行（列）参与变换，而不是仅仅变换其中某些部分，如将下面的所有行加到第一行，或是将第一行加到下面所有行等等，加的时候可能还要乘以一些系数。

（8）（Laplace定理）$\sum_{k=1}^{n} a_{ik} A_{jk} = |\boldsymbol{A}| \delta_{ij} = \begin{cases} |\boldsymbol{A}|, i = j, \\ 0, i \ne j. \end{cases}$（每个元素只有和自己对应的代数余子式相乘，展开式才不是 0）

（9）$|\boldsymbol{AB}| = |\boldsymbol{A}| \cdot |\boldsymbol{B}|$；

（10）对于方阵$\boldsymbol{A},\boldsymbol{B}$，有$\begin{vmatrix}\boldsymbol{A} & * \\ \boldsymbol{O} & \boldsymbol{B}\end{vmatrix}=\begin{vmatrix}\boldsymbol{A} & \boldsymbol{O} \\ * & \boldsymbol{B}\end{vmatrix}=|\boldsymbol{A}|\cdot|\boldsymbol{B}|$.（*表示可以填任意矩阵）

三、逆矩阵

1、逆矩阵的定义：$\boldsymbol{A}^{-1}$满足$\boldsymbol{A}\boldsymbol{A}^{-1}=\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{I}$。

2、逆矩阵存在的充要条件：$|\boldsymbol{A}|\neq 0$。（可逆矩阵也叫非奇异矩阵，不可逆矩阵也叫奇异矩阵）

3、逆矩阵的求解

（1）直接计算：$\boldsymbol{A}^{-1}=\dfrac{\boldsymbol{A}^*}{|\boldsymbol{A}|}$，其中$\boldsymbol{A}^*=(A_{ji})_{n\times n}$（$A_{ji}$是$a_{ji}$的代数余子式）是伴随矩阵。

（注意伴随矩阵的元素下标是ji而不是ij！否则$\boldsymbol{A}$与$\boldsymbol{A}^*$相乘无法让A_{ij}与a_{ij}匹配！）

（2）$(\boldsymbol{A}^{\mathrm{T}})^{-1}=(\boldsymbol{A}^{-1})^{\mathrm{T}}$；

（3）$(\boldsymbol{AB})^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$；

（4）初等变换法：构造矩阵$(\boldsymbol{A}\vdots\boldsymbol{I})$，对其进行初等行变换（换行，倍乘某一行，将某一行乘以一个数加到另一行），直至左边变为$\boldsymbol{I}$，此时右边的矩阵就是$\boldsymbol{A}^{-1}$。

4、n阶可逆方阵行列式的性质：$\left|\boldsymbol{A}^{-1}\right|=\dfrac{1}{|\boldsymbol{A}|},\left|\boldsymbol{A}^*\right|=|\boldsymbol{A}|^{n-1}$.（参考行列式的性质 6）

5、常用分块矩阵公式：$\begin{pmatrix}\boldsymbol{I}_n & \boldsymbol{O} \\ -\boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{I}_m\end{pmatrix}\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D}\end{pmatrix}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\end{pmatrix}$.

推论：若$\boldsymbol{A}$可逆，则$\begin{vmatrix}\boldsymbol{A} & \boldsymbol{D} \\ \boldsymbol{C} & \boldsymbol{B}\end{vmatrix}=|\boldsymbol{A}|\cdot|\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}|$.（必须保证分出的块$\boldsymbol{A}$、$\boldsymbol{D}$是方阵）

6、初等变换矩阵一览

名称	与单位矩阵的差别	作用		逆矩阵
		左乘	右乘	
$\boldsymbol{E}_{ij}$	第ii与jj元素是0； 第ij与ji元素是1	交换第i与j行	交换第i与j列	$\boldsymbol{E}_{ij}$
$\boldsymbol{P}_i(\lambda)$	第i个对角元素变成λ	把第i行倍乘λ	把第i列倍乘λ	$\boldsymbol{P}_i(1/\lambda)$
$\boldsymbol{T}_{ij}(\lambda)$	<u>第ji元素</u>变成λ	把第i行的λ倍加到第j行	把第j列的λ倍加到第i列	$\boldsymbol{T}_{ij}(-\lambda)$

<u>任何可逆矩阵都可以分解成这些初等变换矩阵的乘积。</u>

四、向量的线性关系

1、线性组合：称$\lambda_1\boldsymbol{a}_1+\lambda_2\boldsymbol{a}_2+...+\lambda_n\boldsymbol{a}_n$为向量组$\boldsymbol{a}_1,\boldsymbol{a}_2,...,\boldsymbol{a}_n$的一个线性组合。其中，$\lambda_1,\lambda_2,...,\lambda_n$为实数，称为组合系数。

*线性方程组本质上就是线性组合：令 $\boldsymbol{A} = (\boldsymbol{A}_1,..., \boldsymbol{A}_n)$（列分块），$x = \begin{pmatrix} x_1 \\ ... \\ x_n \end{pmatrix}$，则

$\boldsymbol{Ax} = \boldsymbol{b}$ 等价于 $x_1\boldsymbol{A}_1 + ... + x_n\boldsymbol{A}_n = \boldsymbol{b}$.

2、线性表出：向量 $\boldsymbol{b}$ 可以由向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ 线性表出，是指存在一组实数 $\lambda_1, \lambda_2, ..., \lambda_n$ 使 $\boldsymbol{b} = \lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + ... + \lambda_n\boldsymbol{a}_n$.（**注意：这里的组合系数可以全是 0。**）

3、线性相关的定义：向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ 线性相关，是指存在一组**不全为 0** 的实数 $\lambda_1, \lambda_2, ..., \lambda_n$ 使 $\lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + ... + \lambda_n\boldsymbol{a}_n = \boldsymbol{0}$.（*）

4、线性无关的定义：向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ 线性无关，是指不存在**不全为 0** 的实数 $\lambda_1, \lambda_2, ..., \lambda_n$ 使（*）成立；等价于若要使（*）成立，$\lambda_1, \lambda_2, ..., \lambda_n$ 必全为 0。

5、一个向量线性相关 $\Leftrightarrow$ 此向量为零向量。
两个向量线性相关 $\Leftrightarrow$ 两个向量成比例（共线/平行）。

6、向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ 线性相关（n≥2），等价于其中存在一个向量能够被其他向量线性表出。（再次提醒：线性表出的组合系数可以全是 0）

7、方程组 $\boldsymbol{Ax} = \boldsymbol{b}$ 有无穷多解，等价于其有解且 $\boldsymbol{A}_1,..., \boldsymbol{A}_n$ 线性相关。（注意此定理本身不能说明方程组有解，需事先知道有解）

8、n 个 n 维向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ 线性无关等价于 $|(\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n)| \neq 0$.

9、n 个线性无关 n 维向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ 可以线性表出全部的 n 维向量。

10、若 m 个 n 维向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_m$ 线性相关，则包含 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_m$ 的任意 n 维向量组都线性相关。（部分线性相关则全体线性相关）

11、若 m 维向量 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ 线性无关，$\boldsymbol{b}_1, \boldsymbol{b}_2, ..., \boldsymbol{b}_n$ 是任意的 l 维向量，则两组向量接成的 $m + l$ 维向量 $\begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{b}_1 \end{pmatrix}, \begin{pmatrix} \boldsymbol{a}_2 \\ \boldsymbol{b}_2 \end{pmatrix}, ..., \begin{pmatrix} \boldsymbol{a}_n \\ \boldsymbol{b}_n \end{pmatrix}$ 线性无关。（片段线性无关则整体线性无关）

12、若 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ 线性无关，$\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n, \boldsymbol{b}$ 线性相关，则 $\boldsymbol{b}$ 可由 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ 线性表出，且表达式唯一。

13、子阵：矩阵中任取某些行、列，其交叉点构成的矩阵叫做矩阵的一个子阵。如果取出的子阵是方阵，那么其行列式叫做矩阵的子式。

（1）若矩阵中有一个 r 阶子式不为 0，且所有 $r + 1$ 阶子式全为 0，则此非 0 的 r 阶子式在**原矩阵中**的对应列（行）线性无关，其他列（行）都可用这 r 列（行）

线性表出。
（2）矩阵有 r 列线性无关 $\Leftrightarrow$ 该 r 列中存在非 0 的 r 阶子式。
14、令 $\boldsymbol{A} = (\boldsymbol{A}_1, ..., \boldsymbol{A}_n)$。
$\boldsymbol{A}_1, ..., \boldsymbol{A}_n$ 线性相关 $\Leftrightarrow |\boldsymbol{A}| = 0 \Leftrightarrow$ 存在一组不全为 0 的实数 $\lambda_1, \lambda_2, ..., \lambda_n$ 使（*）成立 $\Leftrightarrow$ 方程组 $\boldsymbol{Ax} = \boldsymbol{0}$ 有非零解，解就是刚才的组合系数。
15、若 $m > n$，则 m 个 n 维向量必然线性相关。

五、向量组与矩阵的秩
1、极大无关组
（1）定义：向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ 中，若存在一个线性无关的子组 $\boldsymbol{a}_{i_1}, \boldsymbol{a}_{i_2}, ..., \boldsymbol{a}_{i_r}$，使得向量组的其他的任何向量都可以由它们线性表出，则称 $\boldsymbol{a}_{i_1}, \boldsymbol{a}_{i_2}, ..., \boldsymbol{a}_{i_r}$ 是该向量组的一个极大无关组。
（2）性质
①任何向量组都有极大无关组。
②同一个向量组可以有不止一个极大无关组，但是它们的元素个数一样。
③向量组中线性无关向量的个数，最多就是极大无关组的向量个数；当向量组中一个线性无关子组的元素个数正好等于极大无关组的向量个数，则它是一个极大无关组。
④扩充定理：向量组 S 中的任何一个线性无关子组都可以被扩充成 S 的一个极大无关组。

2、方阵的迹：$\mathrm{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}$，等于 $\boldsymbol{A}$ 的对角元素之和。
（1）$\mathrm{tr}(\boldsymbol{A} + \boldsymbol{B}) = \mathrm{tr}(\boldsymbol{A}) + \mathrm{tr}(\boldsymbol{B})$。
（2）$\mathrm{tr}(\lambda\boldsymbol{A}) = \lambda\mathrm{tr}(\boldsymbol{A})$。
（3）对于 $m\times n$ 矩阵 $\boldsymbol{A}$ 与 $n\times m$ 矩阵 $\boldsymbol{B}$，$\mathrm{tr}(\boldsymbol{AB}) = \mathrm{tr}(\boldsymbol{BA})$。
（4）如果想要说明两个方阵不相等，可以通过迹不相等来说明。
（例：不存在两个同阶方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 使 $\boldsymbol{AB} - \boldsymbol{BA} = \boldsymbol{I}$）
3、向量组 S 的秩：向量组 S 中极大无关组的向量个数，记为 $\mathrm{rank}(S)$或 $\mathrm{r}(S)$。它反映的是向量组中向量的无关程度、独立程度。
4、若向量组 S 与 T 的向量可以互相线性表出，称 S 与 T 等价。
若 S 中的所有向量都可以由 T 中向量线性表出，则 $\mathrm{r}(S)\leqslant\mathrm{r}(T)$。
若 S 与 T 等价，则 $\mathrm{r}(S) = \mathrm{r}(T)$。
5、矩阵的秩：称矩阵各列（行）构成的向量组的秩为矩阵的列（行）秩。矩阵的行秩与列秩一定相等，称之为矩时阵的秩。
如果 n 阶方阵的秩等于 n，则称该方阵满秩。（**<u>秩的定义不要求是方阵，但是满秩必须是方阵</u>**）
6、秩的性质（**<u>不要求是方阵</u>**）
（1）$\mathrm{r}(\boldsymbol{A} + \boldsymbol{B})\leqslant\mathrm{r}(\boldsymbol{A}) + \mathrm{r}(\boldsymbol{B})$。
（2）若 $\boldsymbol{A}$ 的列数与 $\boldsymbol{B}$ 的行数均为 n，则 $\mathrm{r}(\boldsymbol{A}) + \mathrm{r}(\boldsymbol{B}) - n\leqslant\mathrm{r}(\boldsymbol{AB})\leqslant\min\{\mathrm{r}(\boldsymbol{A}), \mathrm{r}(\boldsymbol{B})\}$。
（3）$\mathrm{r}(\boldsymbol{A}^{\mathrm{T}}) = \mathrm{r}(\boldsymbol{A})$。

（4）$\mathrm{r}(\boldsymbol{O}) = 0$。
（5）若 $k \neq 0$，则 $\mathrm{r}(k\boldsymbol{A}) = \mathrm{r}(\boldsymbol{A})$。
推论：$a,b \neq 0$，则 $\mathrm{r}(a\boldsymbol{A} + b\boldsymbol{B}) \leqslant \mathrm{r}(\boldsymbol{A}) + \mathrm{r}(\boldsymbol{B})$。
（6）初等变换不改变矩阵的秩。
（7）已知 $m \times n$ 矩阵 $\boldsymbol{A}$，且 $\mathrm{r}(\boldsymbol{A}) = r$，则存在 m 阶可逆矩阵 $\boldsymbol{P}$ 与 n 阶可逆矩阵 $\boldsymbol{Q}$ 使 $\boldsymbol{PAQ} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$.
（8）已知 $m \times n$ 矩阵 $\boldsymbol{A}$，对任意 m 阶可逆矩阵 $\boldsymbol{P}$ 与 n 阶可逆矩阵 $\boldsymbol{Q}$，有 $\mathrm{r}(\boldsymbol{PA}) = \mathrm{r}(\boldsymbol{AQ}) = \mathrm{r}(\boldsymbol{A})$。
（9）若 $\boldsymbol{A}$ 为 n 阶方阵，实数 $a \neq b$，则 $(\boldsymbol{A} - a\boldsymbol{I})(\boldsymbol{A} - b\boldsymbol{I}) = \boldsymbol{O}$ 等价于 $\mathrm{r}(\boldsymbol{A} - a\boldsymbol{I}) + \mathrm{r}(\boldsymbol{A} - b\boldsymbol{I}) = n$。
（10）若 $\boldsymbol{A}$ 为 n 阶方阵，则 $\boldsymbol{A}$ 可逆等价于 $\boldsymbol{A}$ 满秩，即 $\mathrm{r}(\boldsymbol{A}) = n$。
7、求矩阵秩的方法
（1）直接根据秩的定义，寻找极大无关组。注意行秩等于列秩，因此秩即不超过行数也不超过列数。如果行数小于列数，则求行秩可能比求列秩更简单，反之亦然。
（2）子式法：矩阵的秩等于其非 0 子式的最大阶数。（寻找矩阵是否存在一个 r 阶子式不为 0 并且 $r + 1$ 阶子式全为 0）
（3）★初等变换构造行阶梯矩阵
①行阶梯矩阵：全 0 行均在最下方，其余行最先出现的非 0 数的列序号从上往下递增。
②每个矩阵（可以不是方阵）经初等变换均可成为行阶梯矩阵。
③行阶梯矩阵的秩等于非 0 行的行数。
④将矩阵化为行阶梯矩阵的规范方法：通过初等变换，使得第 i 列除了前 i 个元素不一定为 0 之外，其他元素全为 0（任意 $1 \leqslant i \leqslant$ 矩阵列数）。这个操作一般按列的顺序来进行，以免混乱。
⑤最简阶梯形：一个矩阵可以变成不同的阶梯形，但是其中的最简阶梯形是唯一的。最简阶梯形要求每个非 0 行的最先出现的非 0 元素就是 1，并且每一列最多只有一个元素不是 0。

六、线性方程组
1、线性方程组的形式
（1）普通形式：$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{cases}$ （未知数与方程个数可能不同）

（2）向量形式：记 $\boldsymbol{A}_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \ldots \\ a_{mi} \end{pmatrix}$，$\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{pmatrix}$，则 $x_1\boldsymbol{A}_1 + \ldots + x_n\boldsymbol{A}_n = \boldsymbol{b}$.

（3）矩阵形式：记 $\boldsymbol{A}=\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & & \\ a_{m1} & \dots & a_{mn} \end{pmatrix}=\left(\boldsymbol{A}_1,\dots,\boldsymbol{A}_n\right)$，则 $\boldsymbol{Ax}=\boldsymbol{b}$。

2、线性方程组解的情况及其判定

（1）线性方程组有解 $\Leftrightarrow$ $\{\boldsymbol{A}_1, ..., \boldsymbol{A}_n\}$ 与 $\{\boldsymbol{A}_1, ..., \boldsymbol{A}_n, \boldsymbol{b}\}$ 等价 $\Leftrightarrow$ $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A}\vdots\boldsymbol{b})$

（2）齐次线性方程组：$\boldsymbol{Ax}=\boldsymbol{0}$；非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}\neq\boldsymbol{0}$。

（3）非齐次线性方程组的通解等于一个特解加上对应齐次方程组的通解。

（4）解的情况：若 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A}\vdots\boldsymbol{b})=n$，则有唯一解；若 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A}\vdots\boldsymbol{b})<n$，则有无穷多解；若 $\mathrm{r}(\boldsymbol{A})\neq\mathrm{r}(\boldsymbol{A}\vdots\boldsymbol{b})$，则无解。（$\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A}\vdots\boldsymbol{b})>n$ 不可能成立，因为矩阵的秩不会超过其列数）

3、线性方程组的解法

（1）Cramer 法则：对于 n 元线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$，如果 $|\boldsymbol{A}|\neq 0$，则方程组有唯一解 $\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}$，且 $\boldsymbol{x}$ 的分量 $x_i=\dfrac{|\boldsymbol{A}_i|}{|\boldsymbol{A}|}$，其中 $|\boldsymbol{A}_i|$ 是将 $|\boldsymbol{A}|$ 的第 i 列换成 $\boldsymbol{b}$ 后的行列式。（计算量略大）

（2）Gauss 消元法：构造增广矩阵 $(\boldsymbol{A}\vdots\boldsymbol{b})$，对其进行初等行变换，直至：

①$\boldsymbol{A}$ 的部分变为三角形矩阵，则可一步步迭代得到结果；

②$\boldsymbol{A}$ 的部分变为单位矩阵，则 $\boldsymbol{b}$ 的部分变为解。

Cramer 法则与 Gauss 消元法一般在解唯一时使用。

（3）基础解系法（适用于一切有解情况）

①判断 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A}\vdots\boldsymbol{b})$ 是否成立——这一步要把增广矩阵 $(\boldsymbol{A}\vdots\boldsymbol{b})$ 化成行阶梯矩阵并求其秩 r，为后续求解做准备。（**<u>处理增广矩阵时不要做列变换，防止改变未知数的顺序</u>**）

②如果 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A}\vdots\boldsymbol{b})$ 成立：设完成变换后的阶梯矩阵为 $(\widetilde{\boldsymbol{A}}\vdots\widetilde{\boldsymbol{b}})=\begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{c} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$，其中 $\boldsymbol{A}_1$ 为 $r\times n$ 矩阵，$\boldsymbol{c}$ 为 r 维列向量。于是原方程组等价于 $\widetilde{\boldsymbol{A}}\boldsymbol{x}=\widetilde{\boldsymbol{b}}$。

③任意设定 n 个未知数中的 r 个为自由变量，用它们表示其他的未知数。不妨设选定的自由变量为 $x_1, x_2, ..., x_r$，且 $\begin{cases} x_{r+1}=\lambda_{11}x_1+\dots+\lambda_{1r}x_r+\mu_1 \\ \dots \\ x_n=\lambda_{n-r,1}x_1+\dots+\lambda_{n-r,r}x_r+\mu_{n-r} \end{cases}$，则有通解

$$\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_r \\ x_{r+1} \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_r \\ \lambda_{11}x_1 + \dots + \lambda_{1r}x_r + \mu_1 \\ \dots \\ \lambda_{n-r,1}x_1 + \dots + \lambda_{n-r,r}x_r + \mu_{n-r} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \\ \mu_1 \\ \dots \\ \mu_{n-r} \end{pmatrix} + x_1\begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \\ \lambda_{11} \\ \dots \\ \lambda_{n-r,1} \end{pmatrix} + x_2\begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \\ \lambda_{12} \\ \dots \\ \lambda_{n-r,2} \end{pmatrix} + \dots + x_r\begin{pmatrix} 0 \\ 0 \\ \dots \\ 1 \\ \lambda_{1r} \\ \dots \\ \lambda_{n-r,r} \end{pmatrix}.$$

其中$\begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \\ \mu_1 \\ \dots \\ \mu_{n-r} \end{pmatrix}$就是一个特解，其余部分$x_1\begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \\ \lambda_{11} \\ \dots \\ \lambda_{n-r,1} \end{pmatrix} + x_2\begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \\ \lambda_{12} \\ \dots \\ \lambda_{n-r,2} \end{pmatrix} + \dots + x_r\begin{pmatrix} 0 \\ 0 \\ \dots \\ 1 \\ \lambda_{1r} \\ \dots \\ \lambda_{n-r,r} \end{pmatrix}$为对应的齐次方程的通解。

一般称$\begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \\ \lambda_{11} \\ \dots \\ \lambda_{n-r,1} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \\ \lambda_{12} \\ \dots \\ \lambda_{n-r,2} \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \dots \\ 1 \\ \lambda_{1r} \\ \dots \\ \lambda_{n-r,r} \end{pmatrix}$为该方程组的一个基础解系，它们是线性无关的。

④如果得到$(\widetilde{\boldsymbol{A}} \vdots \widetilde{\boldsymbol{b}}) = \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{c} \\ \boldsymbol{O} & \boldsymbol{d} \end{pmatrix}$且 $\boldsymbol{d} \neq \boldsymbol{0}$，则原方程组无解。

*（4）Jacobi 迭代法（一般用于数值解，本课程中不常用）：令 $\boldsymbol{A} = \boldsymbol{I} - \boldsymbol{B}$，则 $\boldsymbol{x} = \boldsymbol{B}\boldsymbol{x} + \boldsymbol{C}$。构造向量列$\{\boldsymbol{x}^{(n)}\}$，满足 $\boldsymbol{x}^{(n+1)} = \boldsymbol{B}\boldsymbol{x}^{(n)} + \boldsymbol{C}$（注意右上角加$^{(n)}$表示序号而不是幂次）。选择一个比较合适的初值 $\boldsymbol{x}^{(0)}$，用递推式迭代得到的极限就是原方程组的解。（如果 $\boldsymbol{I} - \boldsymbol{B}$ 这个表达式不好，可以换用其他表达式构造矩阵 $\boldsymbol{B}$，合适即可）

第二章 线性空间、特征值与二次型

一、线性空间的建立与性质

1、常见的线性空间：$\mathbf{R}^n$ 表示 n 维（实）列向量全体；$\mathbf{R}_n$ 表示 n 维（实）行向量全体；$\mathbf{C}^n$ 表示 n 维（复）列向量全体；$C[a,b]$表示$[a,b]$上的连续函数全体；$P_n(x)$ 表示次数不大于 n 的多项式全体。

2、数域：是复数集 $\mathbf{C}$ 的非空非零子集，并且加减乘除运算是封闭的，如实数集

R 与有理数集 **Q**。（一个集合对某运算封闭，是指这个集合中的元素经过此种运算的结果仍在这个集合内。）

3、直积：已知集合 A 与 B，则可定义 A 与 B 的直积 $A\times B = \{(\boldsymbol{x}, \boldsymbol{y})|\boldsymbol{x}\in A, \boldsymbol{y}\in B\}$。一般也记 $A\times A = A^2$。（$\mathbf{R}^n$ 其实就是 n 个 **R** 的直积）

4、线性空间的定义

设 K 为数域，V 为非空集合。如果在 V 中定义加法（$V\times V\mapsto V$）以及数乘（$K\times V\mapsto V$），并且任意 $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\in V$ 与 $a,b\in K$ 均满足以下 8 条性质，则称 V 为 K 上的线性空间。

（1）加法交换律：$\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$

（2）加法结合律：$\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}$

（3）零元存在性：存在 $\mathbf{0}\in V$ 使 $\boldsymbol{x} + \mathbf{0} = \mathbf{0} + \boldsymbol{x} = \boldsymbol{x}$

（4）负元存在性：任意 $\boldsymbol{x}\in V$，存在 $-\boldsymbol{x}\in V$ 使 $\boldsymbol{x} + (-\boldsymbol{x}) = (-\boldsymbol{x}) + \boldsymbol{x} = \mathbf{0}$

（5）数乘幺等律：$1\boldsymbol{x} = \boldsymbol{x}$

（6）数乘左分配律：$(a + b)\boldsymbol{x} = a\boldsymbol{x} + b\boldsymbol{x}$

（7）数乘右分配律：$a(\boldsymbol{x} + \boldsymbol{y}) = a\boldsymbol{x} + a\boldsymbol{y}$

（8）数乘结合律：$(ab)\boldsymbol{x} = a(b\boldsymbol{x})$

如果要证明一个集合是线性空间，只要验证这 8 条性质成立即可。

5、群的定义

在集合 G 中定义乘法“·”，如果满足以下 3 条性质，则称 G 为群。

（1）乘法结合律：$\boldsymbol{x}\cdot(\boldsymbol{y}\cdot\boldsymbol{z}) = (\boldsymbol{x}\cdot\boldsymbol{y})\cdot\boldsymbol{z}$

（2）幺元存在性：存在使 $\boldsymbol{e}\cdot\boldsymbol{x} = \boldsymbol{x}\cdot\boldsymbol{e} = \boldsymbol{x}$

（3）逆元存在性：任意 $\boldsymbol{x}\in G$，存在 $\boldsymbol{x}^{-1}\in G$ 使 $\boldsymbol{x}\cdot\boldsymbol{x}^{-1} = \boldsymbol{x}^{-1}\cdot\boldsymbol{x} = \boldsymbol{e}$

如果还满足第 4 条性质，则称 G 为交换群。

（4）乘法交换律：$\boldsymbol{x}\cdot\boldsymbol{y} = \boldsymbol{y}\cdot\boldsymbol{x}$

*一般线性群：$\mathrm{GL}(K,n)$表示数域 K 上的 n 阶可逆方阵全体。

特殊线性群：$\mathrm{SL}(K,n)$表示数域 K 上的行列式为 1 的 n 阶可逆方阵全体。

6、Hilbert 空间：复数域上一切平方可积函数（可以是复函数），即

$\{f(x)\,|\,x\in\mathbf{C},\int_{-\infty}^{+\infty}|f(x)|^2\mathrm{d}x存在\}$。

7、线性空间的性质

（1）零元与负元的唯一性

（2）数 0 乘任意元素等于零元：$0\boldsymbol{x} = \mathbf{0}$（注意 0 与 $\mathbf{0}$ 的区别）

（3）零元做数乘必等于零元：$a\mathbf{0} = \mathbf{0}$

（4）数乘与负元的关系：$(-1)\boldsymbol{x} = -\boldsymbol{x}$

8、与矩阵相关的线性空间：对于 $m\times n$ 矩阵 $\boldsymbol{A}$，定义列空间/秩空间 $\mathrm{Col}\,\boldsymbol{A} = \{\boldsymbol{A}\boldsymbol{x}|\boldsymbol{x}\in\mathbf{R}^n\}$，以及零空间 $\mathrm{Nul}\,\boldsymbol{A} = \{\boldsymbol{x}\in\mathbf{R}^n|\boldsymbol{A}\boldsymbol{x} = \mathbf{0}\}$。

9、线性子空间

（1）定义：设 V 为线性空间，若 U 是 V 的非空子集，并且 U 按照 V 中的加法与数乘构成线性空间，则 U 是 V 的线性子空间。

（2）线性子空间的判定：U 是 V 的线性子空间 $\Leftrightarrow$ U 关于 V 的加法与数乘在 U 中封闭。（不需要再验证加法与数乘的 8 条性质，因为 V 已经是线性空间）

（3）子空间构造定理：若 V_1 与 V_2 是 V 的两个线性子空间，则 $V_1\cap V_2$ 与 $V_1 + V_2 = \{\boldsymbol{x} + \boldsymbol{y}\,|\,\boldsymbol{x}\in V_1, \boldsymbol{y}\in V_2\}$也是 V 的线性子空间。（注意 $V_1 + V_2$ 不是 $V_1\cup V_2$）

（4）任意线性子空间一定包含零元。

（5）设 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n \in V$，则包含它们的 V 的线性子空间最小为

$\mathrm{span}(\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n) = \{\lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + ... + \lambda_n\boldsymbol{a}_n \mid \lambda_1, \lambda_2, ..., \lambda_n \in K\}$。（作为封闭的线性空间，包含了 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ 就不可避免地包含它们的一切线性组合）

10、即使一个线性空间中的元素不是向量（可以是多项式、矩阵、函数等），也可以按照向量空间中的定义，在这些空间中定义线性相关、线性无关、极大无关组，引入扩充定理与秩。

11、基与维度

（1）基的定义：线性空间 V 中的一个极大无关组称为 V 的一组基。

（2）维数：V 的维数记为 $\dim V$，指 V 中一组基包含的向量个数。若 $\dim V = n$，称 V 为 n 维线性空间。有的线性空间可能是无限维的，例如所有多项式的全体，以及所有的无穷数列全体。

（3）基的扩充定理：V 的任一线性无关子集均可扩充为 V 的一组基。（无论有限维还是无限维均成立）

（4）子空间基的关系：若 V_1 与 V_2 是 V 的两个线性子空间。

设 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_r$ 是 $V_1 \cap V_2$ 的一组基，则它们可扩充为 V_1 的一组基 $\boldsymbol{a}_1, ..., \boldsymbol{a}_r, \boldsymbol{b}_1, ..., \boldsymbol{b}_s$ 和 V_2 的一组基 $\boldsymbol{a}_1, ..., \boldsymbol{a}_r, \boldsymbol{c}_1, ..., \boldsymbol{c}_t$；同时，$\boldsymbol{a}_1, ..., \boldsymbol{a}_r, \boldsymbol{b}_1, ..., \boldsymbol{b}_s, \boldsymbol{c}_1, ..., \boldsymbol{c}_t$ 必为 $V_1 + V_2$ 的一组基。

（5）推论：$\dim V_1 + \dim V_2 = \dim(V_1 \cap V_2) + \dim(V_1 + V_2)$。

二、坐标与坐标变换

1、线性空间 V 的任一元素均可用一组基 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ 以唯一表达式线性表出，也即任意 $\boldsymbol{x} \in V$，存在唯一的一组实数 $\lambda_1, \lambda_2, ..., \lambda_n$ 使得 $\boldsymbol{x} = \lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + ... + \lambda_n\boldsymbol{a}_n$。称 $\boldsymbol{x}$ 在 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ 表示下的坐标向量为 $[\boldsymbol{x}]_{\boldsymbol{a}} = \begin{pmatrix} \lambda_1 \\ ... \\ \lambda_n \end{pmatrix}$.

注：<u>**坐标向量一定是向量**</u>，但是 $\boldsymbol{x}$ 和 $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n$ 可以不是向量。例如，矩阵 $\boldsymbol{x} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$，$\boldsymbol{a}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$，$\boldsymbol{a}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$，$\boldsymbol{a}_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$，$\boldsymbol{a}_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$，则有线性表达式 $\boldsymbol{x} = a\boldsymbol{a}_1 + b\boldsymbol{a}_2 + c\boldsymbol{a}_3 + d\boldsymbol{a}_4$，那么 $[\boldsymbol{x}]_{\boldsymbol{a}} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$. 并不是说，$V$ 中元素是矩阵，坐标就变成矩阵了。

2、设 $\boldsymbol{x}_1, ..., \boldsymbol{x}_n \in V$，则 $\boldsymbol{x}_1, ..., \boldsymbol{x}_n$ 线性相关等价于它们在同一组基表示下的坐标向

量线性相关。

3、坐标变换

（1）同构：称映射$\varphi: \boldsymbol{x} \mapsto [\boldsymbol{x}]_{\boldsymbol{a}}$（$V \mapsto \mathbf{R}^n$）为 V 到 $\mathbf{R}^n$ 的同构。

（2）基变换公式：设$\boldsymbol{a}_1,...,\boldsymbol{a}_n$与$\boldsymbol{b}_1,...,\boldsymbol{b}_n$为 V 的两组基，则有变换公式

$(\boldsymbol{b}_1,...,\boldsymbol{b}_n) = (\boldsymbol{a}_1,...,\boldsymbol{a}_n)\ \left([\boldsymbol{b}_1]_{\boldsymbol{a}},...,[\boldsymbol{b}_n]_{\boldsymbol{a}}\right)$。称$\boldsymbol{A} = \left([\boldsymbol{b}_1]_{\boldsymbol{a}},...,[\boldsymbol{b}_n]_{\boldsymbol{a}}\right)$为从$\boldsymbol{a}_1,...,\boldsymbol{a}_n$到$\boldsymbol{b}_1,...,\boldsymbol{b}_n$的过渡矩阵，则$(\boldsymbol{b}_1,...,\boldsymbol{b}_n) = (\boldsymbol{a}_1,...,\boldsymbol{a}_n)\boldsymbol{A}$。

（3）过渡矩阵必可逆：$(\boldsymbol{a}_1,...,\boldsymbol{a}_n) = (\boldsymbol{b}_1,...,\boldsymbol{b}_n)\ \boldsymbol{A}^{-1}$。与此同时，有

$\left([\boldsymbol{b}_1]_{\boldsymbol{a}},...,[\boldsymbol{b}_n]_{\boldsymbol{a}}\right)^{-1} = \left([\boldsymbol{a}_1]_{\boldsymbol{b}},...,[\boldsymbol{a}_n]_{\boldsymbol{b}}\right)$。（这些坐标向量和基本身一样线性无关）

（4）坐标向量变换公式：$[\boldsymbol{x}]_{\boldsymbol{b}} = \boldsymbol{A}^{-1}[\boldsymbol{x}]_{\boldsymbol{a}}$。

原理：$\boldsymbol{x} = (\boldsymbol{b}_1,...,\boldsymbol{b}_n)[\boldsymbol{x}]_{\boldsymbol{b}} = (\boldsymbol{b}_1,...,\boldsymbol{b}_n)\boldsymbol{A}\boldsymbol{A}^{-1}[\boldsymbol{x}]_{\boldsymbol{b}} = (\boldsymbol{a}_1,...,\boldsymbol{a}_n)[\boldsymbol{x}]_{\boldsymbol{a}}$。

三、线性变换及其矩阵表示

1、线性性：V 上的变换 f 具有线性性，是指任意 $\boldsymbol{x}, \boldsymbol{y} \in V$ 与 $a,b \in K$，有

（1）$f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$；

（2）$f(a\boldsymbol{x}) = af(\boldsymbol{x})$。

由此得出的推论是：$f(a\boldsymbol{x}+b\boldsymbol{y}) = af(\boldsymbol{x}) + bf(\boldsymbol{y})$。

注：①验证一个变换具有线性性，应当验证性质（1）（2）而不是推论。

②常见的具有线性性的变换：极限，定积分，旋转变换，镜面反射，矩阵与向量的乘法。

2、线性变换：已知 $\mathbf{R}$（$\mathbf{C}$）上的线性空间 V 与 U，以及映射$T: V \mapsto U$。若对任意 $\boldsymbol{x}, \boldsymbol{y} \in V$ 与 $a \in \mathbf{R}$（$\mathbf{C}$），有

（1）$T(\boldsymbol{x}+\boldsymbol{y}) = T(\boldsymbol{x}) + T(\boldsymbol{y})$；

（2）$T(a\boldsymbol{x}) = aT(\boldsymbol{x})$。

则称 T 是线性变换/线性映射。

3、线性变换的性质

（1）保零：$T(\mathbf{0}) = \mathbf{0}$

（2）必为奇变换：$T(-\boldsymbol{x}) = -T(\boldsymbol{x})$

（3）$T(a\boldsymbol{x}+b\boldsymbol{y}) = aT(\boldsymbol{x}) + bT(\boldsymbol{y})$

4、线性变换的运算

（1）加映射：$(T_1+T_2)(\boldsymbol{x}) = T_1(\boldsymbol{x}) + T_2(\boldsymbol{x})$

（2）数乘映射：$(aT)(\boldsymbol{x}) = aT(\boldsymbol{x})$

（3）复合映射：$(T_2 \circ T_1)(\boldsymbol{x}) = T_2[T_1(\boldsymbol{x})]$

（4）逆映射：若对任意 $\boldsymbol{y} \in U$，存在唯一的 $\boldsymbol{x} \in V$ 使得 $T(\boldsymbol{x}) = \boldsymbol{y}$，则存在逆映射 $T^{-1}: U \mapsto V$ 使得 $T^{-1}(\boldsymbol{y}) = \boldsymbol{x}$。

①逆映射与原映射的关系：任意 $\boldsymbol{y} \in U$，$T[T^{-1}(\boldsymbol{y})] = \boldsymbol{y}$；任意 $\boldsymbol{x} \in V$，$T^{-1}[T(\boldsymbol{x})] = \boldsymbol{x}$。

*②$V \mapsto V$ 的线性映射 T 为可逆映射$\Leftrightarrow$ T 是单射$\Leftrightarrow$ T 是满射$\Leftrightarrow$ T 是一一映射。

5、线性变换的表示矩阵：任意线性变换 T 都有一个表示矩阵 $\boldsymbol{A}$；对一个向量做线性变换相当于该向量左乘表示矩阵，即 $T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$。

（1）已知线性变换 $T:\mathbf{R}^n \mapsto \mathbf{R}^m$。若用 $\mathbf{R}^n$ 的基 $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n$ 表示 $\boldsymbol{x}$，用 $\mathbf{R}^m$ 的基 $\boldsymbol{b}_1,\ldots,\boldsymbol{b}_m$ 表示 $T(\boldsymbol{x})$，则有 $\boldsymbol{A} = \left([T(\boldsymbol{a}_1)]_{\boldsymbol{b}},\ldots,[T(\boldsymbol{a}_n)]_{\boldsymbol{b}}\right)$.

（2）上述过程中，若 $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n$ 是 $\mathbf{R}^n$ 的标准基，即 $\boldsymbol{a}_i$ 的第 i 个分量是 1 而其他分量都是 0（$1\leqslant i\leqslant n$），以及 $\boldsymbol{b}_1,\ldots,\boldsymbol{b}_m$ 是 $\mathbf{R}^m$ 的标准基则 $\boldsymbol{A} = \left(T(\boldsymbol{a}_1),\ldots,T(\boldsymbol{a}_n)\right)$.（在标准基下，向量的坐标向量等于它自身）

（3）设线性映射 T_1 与 T_2 的表示矩阵分别是 $\boldsymbol{A}$ 与 $\boldsymbol{B}$，则 $T_1 + T_2$ 的表示矩阵为 $\boldsymbol{A}+\boldsymbol{B}$，$aT_1$ 的表示矩阵为 $a\boldsymbol{A}$，$T_2\circ T_1$ 的表示矩阵为 $\boldsymbol{B}\boldsymbol{A}$。

（4）设可逆线性映射 T 的表示矩阵为 $\boldsymbol{A}$，则 T^{-1} 的表示矩阵为 $\boldsymbol{A}^{-1}$。

（5）一个重要的表示矩阵：在二维空间中，将一个向量逆时针旋转θ角，对应旋转矩阵 $\boldsymbol{A}(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$；其逆变换相当于顺时针旋转$\theta$，也就是逆时针旋转-$\theta$，故 $\boldsymbol{A}^{-1}(\theta) = \boldsymbol{A}(-\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$。

6、矩阵的相似性

（1）线性变换的表示矩阵与选用的基有关，因为向量的坐标总是依赖于基。（概念辨析：不同的基只会影响向量的坐标，但是这不代表这个向量发生变化；不同坐标只是**<u>对同一主体的不同描述方式</u>**，不要把它看成向量本身的绝对属性。）

（2）已知线性变换 $T:V \mapsto V$。设 $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n$ 与 $\boldsymbol{b}_1,\ldots,\boldsymbol{b}_n$ 是 V 的两组基，T 在这两组基下的表示矩阵分别是 $\boldsymbol{A}$ 与 $\boldsymbol{B}$，从 $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n$ 到 $\boldsymbol{b}_1,\ldots,\boldsymbol{b}_n$ 的过渡矩阵为 $\boldsymbol{P}$，则有

①$[\boldsymbol{x}]_{\boldsymbol{b}} = \boldsymbol{P}^{-1}[\boldsymbol{x}]_{\boldsymbol{a}}$；②$[T(\boldsymbol{x})]_{\boldsymbol{b}} = \boldsymbol{P}^{-1}[T(\boldsymbol{x})]_{\boldsymbol{a}}$；③$[T(\boldsymbol{x})]_{\boldsymbol{a}} = \boldsymbol{A}[\boldsymbol{x}]_{\boldsymbol{a}}$；④$[T(\boldsymbol{x})]_{\boldsymbol{b}} = \boldsymbol{B}[\boldsymbol{x}]_{\boldsymbol{b}}$。

综合四个关系，$[T(\boldsymbol{x})]_{\boldsymbol{b}} = \boldsymbol{P}^{-1}[T(\boldsymbol{x})]_{\boldsymbol{a}} = \boldsymbol{P}^{-1}\boldsymbol{A}[\boldsymbol{x}]_{\boldsymbol{a}} = \boldsymbol{B}[\boldsymbol{x}]_{\boldsymbol{b}} = \boldsymbol{B}\boldsymbol{P}^{-1}[\boldsymbol{x}]_{\boldsymbol{a}}$。

$\boldsymbol{P}^{-1}\boldsymbol{A}[\boldsymbol{x}]_{\boldsymbol{a}} = \boldsymbol{B}\boldsymbol{P}^{-1}[\boldsymbol{x}]_{\boldsymbol{a}} \Rightarrow$ 不同基下表示矩阵的关系：$\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$（注意每个矩阵的含义）

（3）相似的定义：已知 n 阶方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$，若存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$，则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似，记为 $\boldsymbol{A}\sim\boldsymbol{B}$。（内涵：两个矩阵能在不同基下表示相同的线性变换）

（4）相似矩阵的性质

①自反性：$\boldsymbol{A}\sim\boldsymbol{A}$

②对称性：若 $\boldsymbol{A}\sim\boldsymbol{B}$，则 $\boldsymbol{B}\sim\boldsymbol{A}$。

③传递性：若 $\boldsymbol{A}\sim\boldsymbol{B}$，$\boldsymbol{B}\sim\boldsymbol{C}$，则 $\boldsymbol{A}\sim\boldsymbol{C}$。

④幂次：若 $\boldsymbol{A}\sim\boldsymbol{B}$，$k$ 为正整数，则 $\boldsymbol{A}^k\sim\boldsymbol{B}^k$。

7、与线性变换相关的空间

（1）核空间：$\text{Ker } T=\{\boldsymbol{x}\in V\mid T(\boldsymbol{x})=\boldsymbol{0}\}$（$T$ 是单射 $\Rightarrow \text{Ker } T=\{\boldsymbol{0}\}\Leftrightarrow \dim \text{Ker } T=0$）

（2）像空间：$\text{Im } T=\{T(\boldsymbol{x})\mid \boldsymbol{x}\in V\}$

（3）维度公式：若 $T:V\mapsto U$ 在 V 和 U 的某基下的表示矩阵为 $\boldsymbol{A}_{m\times n}$，则有 $\dim \text{Im } T=\text{r}(\boldsymbol{A})$，$\dim \text{Ker } T=n-\text{r}(\boldsymbol{A})$；$\dim \text{Im } T+\dim \text{Ker } T=\dim V$。

四、特征值

1、特征值的定义：已知 n 阶方阵 $\boldsymbol{A}$。若存在 $\lambda\in\mathbf{C}$，以及非零向量 $\boldsymbol{x}\in\mathbf{C}^n$，使得 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$，则称为 λ 为 $\boldsymbol{A}$ 的一个特征值，$\boldsymbol{x}$ 为 $\boldsymbol{A}$ 的对应于 λ 的一个特征向量。如果 $\boldsymbol{A}$ 对应线性变换 T，则 $T(\boldsymbol{x})=\lambda\boldsymbol{x}$，也可定义线性变换 T 的特征值与特征向量。

2、特征值的求解：$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}\Leftrightarrow(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$ 有非零解 $\Leftrightarrow|\boldsymbol{A}-\lambda\boldsymbol{I}|=0$。

3、特征多项式：定义矩阵 $\boldsymbol{A}$ 的特征多项式 $f(\lambda)=|\boldsymbol{A}-\lambda\boldsymbol{I}|$。

（1）主子式：在一个矩阵中，选取序号相同的 k 个行与列的交叉点（例如第 1、3 行与第 1、3 列的交叉点），构成的 k 阶行列式为该矩阵的一个 k 阶主子式。

（2）用主子式求特征多项式：记 n 阶方阵 $\boldsymbol{A}$ 的所有 k 阶主子式之和为 S_k，则 $\left|\boldsymbol{A}-\lambda\boldsymbol{I}\right|=(-1)^n\sum_{k=0}^{n}(-1)^k S_k\lambda^{n-k}$. 其中，规定 $S_0=1$。（用主子式求特征多项式，在大多数情况下比直接化简 $|\boldsymbol{A}-\lambda\boldsymbol{I}|$ 更加方便，而且出错几率更小。）

（3）一些主子式可以快速判断：$S_0=1$，$S_1=\sum_{i=1}^{n}a_{ii}=\text{tr}(\boldsymbol{A})$，$S_n=|\boldsymbol{A}|$。

4、特征值的性质

（1）n 阶方阵 $\boldsymbol{A}$ 必然在复数域上有 n 个特征值 $\lambda_1,\lambda_2,\ldots,\lambda_n$（重根不合并）。

（2）韦达定理：$\lambda_1+\lambda_2+\ldots+\lambda_n=\text{tr}(\boldsymbol{A})=\sum_{i=1}^{n}a_{ii}$；$\lambda_1\lambda_2\ldots\lambda_n=|\boldsymbol{A}|$。

（3）若 λ 是 $\boldsymbol{A}$ 的特征值，且 $\boldsymbol{A}$ 可逆，则 $1/\lambda$ 必为 $\boldsymbol{A}^{-1}$ 的特征值。

（4）若 λ 是 $\boldsymbol{A}$ 的特征值，则 $k_0+k_1\lambda+\ldots+k_m\lambda^m$ 必为 $k_0\boldsymbol{I}+k_1\boldsymbol{A}+\ldots+k_m\boldsymbol{A}^m$ 的特征值。

（5）若 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$，则对于任意正整数 k，有 $\boldsymbol{A}^k\boldsymbol{x}=\lambda^k\boldsymbol{x}$。

（6）一个矩阵的互异的两个特征值对应的特征向量线性无关；同一个非重根的特征值对应的线性无关特征向量只有一个，重根不一定。

5、特征值与特征向量的意义

（1）线性变换的特征向量，就是经过该变换只变长度、不变方向的向量；对应的特征值就是把这个向量放大的倍数。

★（2）用特征向量计算线性变换 $\boldsymbol{A}^k\boldsymbol{x}$ 的结果

①求 $\boldsymbol{A}$ 的特征值 $\lambda_1,\lambda_2,\ldots,\lambda_n$，分别对应特征向量 $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n$。

②用以上特征向量将 $\boldsymbol{x}$ 线性表出：$\boldsymbol{x}=\mu_1\boldsymbol{a}_1+\ldots+\mu_n\boldsymbol{a}_n$。

③计算结果：$\boldsymbol{A}^k\boldsymbol{x}=\boldsymbol{A}^k(\mu_1\boldsymbol{a}_1+\ldots+\mu_n\boldsymbol{a}_n)=\mu_1\lambda_1^k\boldsymbol{a}_1+\ldots+\mu_n\lambda_n^k\boldsymbol{a}_n$。

（特征向量的作用：将任意的向量分解为特征向量的组合，则可以很大程度上简化 $\boldsymbol{A}^k\boldsymbol{x}$ 的计算，完全不用做矩阵乘法，只需把原先的各个特征向量乘以某个系数。）

五、矩阵的对角化

1、可对角化：对于 n 阶方阵 $\boldsymbol{A}$，若存在可逆矩阵 $\boldsymbol{P}$ 使

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\lambda_1,\lambda_2,...,\lambda_n)=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & ... & \\ & & & \lambda_n\end{pmatrix}$$，则称 $\boldsymbol{A}$ 可对角化。

2、相似矩阵有相同特征值：

$$\left|\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}-\lambda\boldsymbol{I}\right|=\left|\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}-\lambda\boldsymbol{P}^{-1}\boldsymbol{I}\boldsymbol{P}\right|=\left|\boldsymbol{P}^{-1}\boldsymbol{P}\right|\left|\boldsymbol{A}-\lambda\boldsymbol{I}\right|=\left|\boldsymbol{A}-\lambda\boldsymbol{I}\right|.$$

因此。对角化得到的对角矩阵$\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & ... & \\ & & & \lambda_n\end{pmatrix}$，其对角元素就是 $\boldsymbol{A}$ 的特征值。

3、n 阶方阵 $\boldsymbol{A}$ 可对角化的充要条件

（1）$\boldsymbol{A}$ 可对角化 $\Leftrightarrow$ $\boldsymbol{A}$ 有 n 个线性无关的特征向量。（充分不必要条件：$\boldsymbol{A}$ 有 n 个不同的特征值）

（2）设 $\boldsymbol{A}$ 有 k 个互异特征值 $\lambda_1,\lambda_2,...,\lambda_k$，其中$\lambda_i$ 的重数为 n_i（$1\leqslant i\leqslant k$）。

任意方阵：$\mathrm{r}(\boldsymbol{A}-\lambda_i\boldsymbol{I})\geqslant n-n_i$；$\boldsymbol{A}$ 可对角化 $\Leftrightarrow$ $\mathrm{r}(\boldsymbol{A}-\lambda_i\boldsymbol{I})=n-n_i$（$1\leqslant i\leqslant k$）$\Leftrightarrow$ λ_i 有 n_i 个线性无关的特征向量（$1\leqslant i\leqslant k$）。

4、对角矩阵的性质（简化计算）：$\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & ... & \\ & & & \lambda_n\end{pmatrix}^m=\begin{pmatrix}\lambda_1^m & & & \\ & \lambda_2^m & & \\ & & ... & \\ & & & \lambda_n^m\end{pmatrix}$。

5、**变换矩阵$\boldsymbol{P}$的寻找：若**$\boldsymbol{a}_1,...,\boldsymbol{a}_n$ 分别对应特征值 $\lambda_1,\lambda_2,...,\lambda_n$，则 $\boldsymbol{P}=(\boldsymbol{a}_1,...,\boldsymbol{a}_n)$ 可使得$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & ... & \\ & & & \lambda_n\end{pmatrix}$。（注意 $\boldsymbol{P}$ 的每一列和对角阵中的元素是对应的，顺序不可以乱）

6、Jordan 标准型

（1）定义：m 阶上三角阵 $\boldsymbol{J}_m(x)=\begin{pmatrix} x & 1 & & \\ & x & \ldots & \\ & & \ldots & 1 \\ & & & x \end{pmatrix}$

（2）若 n 阶复方阵 $\boldsymbol{A}$ 有 k 个互异特征值 $\lambda_1,\lambda_2,\ldots,\lambda_k$，$\lambda_i$ 的重数为 n_i（$1\leqslant i\leqslant k$），

则 $\boldsymbol{A}$ 必定相似于上三角阵 $\boldsymbol{J}=\begin{pmatrix} \boldsymbol{J}_{n_1}(\lambda_1) & & & \\ & \boldsymbol{J}_{n_2}(\lambda_2) & & \\ & & \ldots & \\ & & & \boldsymbol{J}_{n_k}(\lambda_k) \end{pmatrix}$。

（3）$\boldsymbol{J}_m(x)$的性质

①分解：$\boldsymbol{J}_m(x)=x\boldsymbol{I}_m+\boldsymbol{J}_m(0)$。

②$\boldsymbol{J}_m(0)\begin{pmatrix} a_{11} & \ldots & a_{1m} \\ a_{21} & \ldots & a_{2m} \\ \ldots & & \\ a_{m-1,1} & \ldots & a_{m-1,m} \\ a_{m1} & \ldots & a_{mm} \end{pmatrix}=\begin{pmatrix} a_{21} & \ldots & a_{2m} \\ a_{31} & \ldots & a_{3m} \\ \ldots & & \\ a_{m1} & \ldots & a_{mm} \\ 0 & \ldots & 0 \end{pmatrix}$。（左乘 $\boldsymbol{J}_m(0)$的作用："吞去"第一行，其他行上移，最后一行变成 0）

推论：$[\boldsymbol{J}_m(0)]^m=\boldsymbol{O}$。

③$[\boldsymbol{J}_m(x)]^k=[x\boldsymbol{I}_m+\boldsymbol{J}_m(0)]^k=\sum_{i=0}^{k}C_k^i x^{k-i}[\boldsymbol{J}_m(0)]^i$。（后面一些项可能是 0）

最终结果：$[\boldsymbol{J}_m(x)]^k=\begin{pmatrix} x^k & kx^{k-1} & \frac{k(k-1)}{2}x^{k-2} & \ldots & \\ & x^k & kx^{k-1} & \ldots & \\ & & x^k & \ldots & \\ & & & \ldots & \\ & & & & x^k \end{pmatrix}_{m\times m}$。

六、内积空间

1、实空间中的内积：若 $\boldsymbol{x},\boldsymbol{y}\in\mathbf{R}^n$，则有内积$(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}=(y_1,\ldots,y_n)\begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix}$.

2、内积的性质

（1）对称性：$(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{y},\boldsymbol{x})$

（2）左侧线性：$(a\boldsymbol{x}+b\boldsymbol{y},\boldsymbol{z})=a(\boldsymbol{x},\boldsymbol{z})+b(\boldsymbol{y},\boldsymbol{z})$（同理有右侧线性）

（3）正定性：$(\boldsymbol{x},\boldsymbol{x})\geqslant 0$，当且仅当 $\boldsymbol{x}=\boldsymbol{0}$ 时取等号。

（4）Cauchy-Schwarz 不等式：$(\boldsymbol{x},\boldsymbol{y})^2\leqslant(\boldsymbol{x},\boldsymbol{x})(\boldsymbol{y},\boldsymbol{y})$

（5）三角不等式：$\|\boldsymbol{x}+\boldsymbol{y}\|\leqslant\|\boldsymbol{x}\|+\|\boldsymbol{y}\|$

3、Euclid 空间：定义了内积的线性空间。

4、长度与夹角

（1）长度/范数：$\|\boldsymbol{x}\|=\sqrt{(\boldsymbol{x},\boldsymbol{x})}$. 当$\|\boldsymbol{x}\|=1$时，称 $\boldsymbol{x}$ 为单位向量（这个“向量”是广义的线性空间的元素，不是狭义的向量）。

（2）夹角：$\langle\boldsymbol{x},\boldsymbol{y}\rangle=\arccos\dfrac{(\boldsymbol{x},\boldsymbol{y})}{\|\boldsymbol{x}\|\cdot\|\boldsymbol{y}\|}$.（范围为$[0,\pi]$）。如果 $\boldsymbol{x}=\boldsymbol{0}$ 或 $\boldsymbol{y}=\boldsymbol{0}$，夹角大小任取。

（3）垂直/正交：当$(\boldsymbol{x},\boldsymbol{y})=0$ 时，$\langle\boldsymbol{x},\boldsymbol{y}\rangle=\dfrac{\pi}{2}$，此时称 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 垂直/正交。

5、正交基与正交化

（1）若线性空间的一组基两两正交，则称它们为一组正交基；若这些正交基的长度均为 1，则称它们为一组标准正交基。

（2）线性无关是正交的必要不充分条件。

（3）对于一组正交向量$\boldsymbol{a}_1,...,\boldsymbol{a}_n$，若$\boldsymbol{b}=\lambda_1\boldsymbol{a}_1+...+\lambda_n\boldsymbol{a}_n$，则$\lambda_i=\dfrac{(\boldsymbol{b},\boldsymbol{a}_i)}{(\boldsymbol{a}_i,\boldsymbol{a}_i)}$.（$1\leqslant i\leqslant n$）

以上推导要求$\boldsymbol{b}\in\mathrm{span}(\boldsymbol{a}_1,...,\boldsymbol{a}_n)\equiv S$；若$\boldsymbol{b}\notin S$，则必有另一个向量

$\widetilde{\boldsymbol{b}}=\sum\limits_{i=1}^{n}\dfrac{(\boldsymbol{b},\boldsymbol{a}_i)}{(\boldsymbol{a}_i,\boldsymbol{a}_i)}\boldsymbol{a}_i\in S$。称$\widetilde{\boldsymbol{b}}=\mathrm{Proj}_S\boldsymbol{b}$为 $\boldsymbol{b}$ 在空间 S 中的正投影，它满足

$(\boldsymbol{b}-\widetilde{\boldsymbol{b}},\boldsymbol{a}_i)=0$。（$1\leqslant i\leqslant n$）

（4）Schmidt 正交化：将一组线性无关但是不正交的基转化成正交基。

已知线性无关向量$\boldsymbol{a}_1,...,\boldsymbol{a}_n$。令

$\boldsymbol{b}_1=\boldsymbol{a}_1$，$i\geqslant 2$ 时$\boldsymbol{b}_i=\boldsymbol{a}_i-\sum\limits_{j=1}^{i-1}\dfrac{(\boldsymbol{a}_i,\boldsymbol{b}_j)}{(\boldsymbol{b}_j,\boldsymbol{b}_j)}\boldsymbol{b}_j=\boldsymbol{a}_i-\mathrm{Proj}_{\mathrm{span}(\boldsymbol{b}_1,...,\boldsymbol{b}_{i-1})}\boldsymbol{a}_i$，则$\boldsymbol{b}_1,...,\boldsymbol{b}_n$是与$\boldsymbol{a}_1,...,\boldsymbol{a}_n$等价的一组正交向量。将这些向量全部单位化，可以得到一组标准正交向量。

（5）（标准）正交基扩充定理：$\mathbf{R}^n$ 中任意一组（标准）正交向量可以扩充为 $\mathbf{R}^n$ 的一组（标准）正交基。

6、正交矩阵

（1）标准正交基$\boldsymbol{e}_1,...,\boldsymbol{e}_n$的性质：$(\boldsymbol{e}_i,\boldsymbol{e}_j)=\delta_{ij}=\begin{cases}1,i=j\\0,i\neq j\end{cases}$

（2）正交矩阵的定义：若 n 阶方阵 $\boldsymbol{Q}$ 的各列为 $\mathbf{R}^n$ 的**标准**正交基，则称 $\boldsymbol{Q}$ 为正交矩阵。

（3）正交矩阵 $\boldsymbol{Q}$ 的性质

①$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q}=\boldsymbol{I}$，即 $\boldsymbol{Q}^{\mathrm{T}}=\boldsymbol{Q}^{-1}$（是 $\boldsymbol{Q}$ 为正交阵的充要条件）；

②$|\boldsymbol{Q}|=\pm1$；

③$\boldsymbol{Q}$ 的特征值（可以是复数）的模长都是 1；
④$\boldsymbol{Q}^{\mathrm{T}} = \boldsymbol{Q}^{-1}$ 也是正交阵；
⑤两个正交阵的乘积仍为正交阵。
（4）同一线性空间的任意两组标准正交基之间的过渡矩阵为正交阵。

7、正交变换

（1）定义：设线性变换 $T:\mathbf{R}^n \mapsto \mathbf{R}^n$。若对任意 $\boldsymbol{x}, \boldsymbol{y}\in\mathbf{R}^n$，$(T(\boldsymbol{x}), T(\boldsymbol{y})) = (\boldsymbol{x}, \boldsymbol{y})$，则称 T 为正交变换/保积变换。
（2）T 为正交变换的充要条件：
①T 保持长度，即$\|T(\boldsymbol{x})\| = \|\boldsymbol{x}\|$；
②T 将 $\mathbf{R}^n$ 中的任意一组标准正交基映射成标准正交基；
③T 将 $\mathbf{R}^n$ 中的某一组标准正交基映射成标准正交基；
④T 在标准正交基下的表示矩阵为正交阵。
（以上所有命题之间是等价的；②③不是重复命题，因为它暗含对某一组标准正交基则对任意一组标准正交基成立）
（3）已知 $\mathbf{R}^n$ 的一组标准正交基 $\boldsymbol{e}_1,\dots,\boldsymbol{e}_n$，则有$(\boldsymbol{x}, \boldsymbol{y}) = ([\boldsymbol{x}]_{\boldsymbol{e}}, [\boldsymbol{y}]_{\boldsymbol{e}})$。

8、复空间的内积

（1）定义：若 $\boldsymbol{x}, \boldsymbol{y}\in\mathbf{C}^n$，则有内积$(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{y}^{\mathrm{H}}\boldsymbol{x} = (\bar{y}_1,\dots,\bar{y}_n)\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}$.

（2）性质

①共轭对称性：$(\boldsymbol{x}, \boldsymbol{y}) = \overline{(\boldsymbol{y}, \boldsymbol{x})}$

②左侧线性：$(a\boldsymbol{x} + b\boldsymbol{y}, \boldsymbol{z}) = a(\boldsymbol{x}, \boldsymbol{z}) + b(\boldsymbol{y}, \boldsymbol{z})$（$a, b\in\mathbf{C}$）

③右侧共轭线性：$(\boldsymbol{z}, a\boldsymbol{x} + b\boldsymbol{y}) = \bar{a}\overline{(\boldsymbol{x}, \boldsymbol{z})} + \bar{b}\overline{(\boldsymbol{y}, \boldsymbol{z})}$（不是右侧线性!!）

④正定性：$(\boldsymbol{x}, \boldsymbol{x})$必为实数，并且$(\boldsymbol{x}, \boldsymbol{x})\geqslant 0$；当且仅当 $\boldsymbol{x} = \boldsymbol{0}$ 时取等号。
（实际上，$(\boldsymbol{x}, \boldsymbol{x}) = \|\boldsymbol{x}\|^2 = |x_1|^2 +\dots+|x_n|^2$）
⑤Cauchy-Schwarz 不等式：$|(\boldsymbol{x}, \boldsymbol{y})|^2 \leqslant (\boldsymbol{x}, \boldsymbol{x})(\boldsymbol{y}, \boldsymbol{y})$
⑥三角不等式：$\|\boldsymbol{x} + \boldsymbol{y}\|\leqslant\|\boldsymbol{x}\| + \|\boldsymbol{y}\|$

9、酉矩阵

（1）定义：若 n 阶复方阵 $\boldsymbol{U}$ 的各列在 $\mathbf{C}^n$ 定义的内积下是标准正交基（相互正交，范数均为 1），则称 $\boldsymbol{U}$ 为酉矩阵。
（2）酉矩阵的性质
①$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U} = \boldsymbol{I}$，即 $\boldsymbol{U}^{\mathrm{H}} = \boldsymbol{U}^{-1}$；
②特征值的模长均为 1；
③$\boldsymbol{U}^{\mathrm{H}} = \boldsymbol{U}^{-1}$ 也是酉矩阵；
④两个酉矩阵的乘积仍为酉矩阵。

10、广义的内积定义

（1）任意实线性空间上定义的运算，只要满足对称性、左侧线性、正定性，就可以说该运算是实空间上的内积。
（2）任意复线性空间上定义的运算，只要满足共轭对称性、左侧线性、正定性，

就可以说该运算是复空间上的内积。
（3）验证一个抽象空间的运算是否为内积，就是验证对应的三条性质是否成立。

七、正交相似与酉相似
1、正交相似的定义：已知 n 阶实方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$，若存在正交矩阵 $\boldsymbol{P}$ 使得

$\boldsymbol{B}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$（$=\boldsymbol{P}^{-1}\boldsymbol{AP}$），则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 正交相似。

2、酉相似的定义：已知 n 阶复方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$，若存在酉矩阵 $\boldsymbol{U}$ 使得

$\boldsymbol{B}=\boldsymbol{U}^{\mathrm{H}}\boldsymbol{AU}$（$=\boldsymbol{U}^{-1}\boldsymbol{AU}$），则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 酉相似。（正交矩阵就是实的酉矩阵。）

3、Schur 定理：任一实方阵正交相似于实的上三角阵；任一复方阵酉相似于复的上三角阵。
4、实对称阵与 Hermite 阵
（1）定义：若 n 阶实方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}$，则称 $\boldsymbol{A}$ 为实对称阵。
若 n 阶复方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{H}}$，则称 $\boldsymbol{A}$ 为 Hermite 阵。
（2）实对称阵与 Hermite 阵的性质
①特征值均为实数；
②不同特征值对应的特征向量均正交；
③有 n 个相互正交的特征向量。
5、实方阵 $\boldsymbol{A}$ **正交相似**于**实对角阵**的充要条件是 $\boldsymbol{A}$ 为实对称阵，即 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}$。
复方阵 $\boldsymbol{A}$ **酉相似**于**实对角阵**的充要条件是 $\boldsymbol{A}$ 为 Hermite 阵，即 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{H}}$。
复方阵 $\boldsymbol{A}$ **酉相似**于**复对角阵**的充要条件是 $\boldsymbol{A}$ 为正规矩阵，即 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$。
对比：方阵 $\boldsymbol{A}$ 相似于对角阵的充要条件是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量，或者 $\mathrm{r}(\boldsymbol{A}-\lambda_i\boldsymbol{I})=n-n_i$（$1\leqslant i\leqslant k$）。

八、二次型问题
1、二次型的定义与书写
（1）文字定义：二次型指实系数二次齐次多项式。

（2）代数式写法：$\sum_{i,j=1}^{n}a_{ij}x_ix_j=a_{11}x_1^2+a_{12}x_1x_2+\dots+a_{nn}x_n^2$。

（3）线性代数写法：令 $\boldsymbol{x}=\begin{pmatrix}x_1\\ \dots\\ x_n\end{pmatrix}$，$\boldsymbol{A}=\left(a_{ij}\right)_{n\times n}$，则上述二次型可以写为 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}$。

2、相伴矩阵：由于交叉项 x_ix_j 与 x_jx_i 是可以合并的，所以我们默认二者的系数平均分配（相等）。这样得到的矩阵 $\boldsymbol{A}$ 是实对称阵，称为该二次型的相伴矩阵。（暗示着相伴矩阵一定是实对称阵）
3、不同变量表示下二次型的相伴矩阵：若令 $\boldsymbol{x}=\boldsymbol{Py}$，则 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{APy}$。记 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}=\boldsymbol{B}$，则 $\boldsymbol{B}$ 是新的相伴矩阵。
4、矩阵的合同性
（1）定义：已知 n 阶实对称阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$，若存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{B}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$，则

称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同。（这里 $\boldsymbol{P}$ 不一定是正交阵，所以这个定义和正交相似不一样）
（2）性质
①自反性：实对称阵与自身必合同。
②对称性：若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同，则 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 合同。
③传递性：若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同，$\boldsymbol{B}$ 与 $\boldsymbol{C}$ 合同，则 $\boldsymbol{A}$ 与 $\boldsymbol{C}$ 合同。
5、标准型：通过坐标变换可以将二次型转化成无交叉项的形式，这个形式叫做标准型；如果进一步变换，使得每一项的系数都是 0 或 ± 1，对应的形式叫做规范型。任何二次型都有标准型与规范型；标准型不唯一，但是规范型一定唯一。（这里不考虑系数顺序，只是说规范型中不同系数的值与出现的次数是固定的）
★6、将二次型化为标准型的几种方法
（1）将二次型化为标准型，实质就是通过合同变换将相伴矩阵化成对角阵。
（2）配方法：非常基础，但是对于多变量情况有些难处理，在此从略。
（3）正交相似法：对于正交矩阵 $\boldsymbol{P}$，合同变换 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$ 与相似变换 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 没有区别。此时的做法和以前的对角化完全一样：
①求出 $\boldsymbol{A}$ 的特征值和 n 个相互正交的特征向量（$\boldsymbol{A}$ 是实对称阵，因此可以保证它们正交）；
②将这些特征向量单位化，排成矩阵 $\boldsymbol{P}$，则 $\boldsymbol{P}$ 是正交阵；
③求出 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$，得到对角阵。
（正交相似法在对角化方面较方便，但是如果必须求出变换矩阵 $\boldsymbol{P}$ 则比较麻烦）
（4）初等变换法：任何可逆矩阵都可以分解成这些初等变换矩阵的乘积。

名称	作用	
	左乘 $\boldsymbol{E}_{ij}/\boldsymbol{P}_i(\lambda)/\boldsymbol{T}_{ij}(\lambda)$	右乘 $\boldsymbol{E}_{ij}^{\mathrm{T}}/\boldsymbol{P}_i(\lambda)^{\mathrm{T}}/\boldsymbol{T}_{ij}(\lambda)^{\mathrm{T}}$
$\boldsymbol{E}_{ij}$	交换第 i 与 j 行	交换第 i 与 j 列
$\boldsymbol{P}_i(\lambda)$	把第 i 行倍乘λ	把第 i 列倍乘λ
$\boldsymbol{T}_{ij}(\lambda)$	把第 i 行的λ倍加到第 j 行	把第 i 列的λ倍加到第 j 列

①由于这里要做的是合同变换，每做一次行变换，就必须做对应的列变换；“对应”是指上表中左乘某矩阵和右乘它的转置相对应。
②如果可以直接通过行变换将 $\boldsymbol{A}$ 化为上三角阵，则对应的列变换会继续将它化成对角阵。因此，只要知道了每次做的行变换对应的初等矩阵，将它们按先后顺序<u>从右往左</u>乘起来，得到的矩阵就是 $\boldsymbol{P}^{\mathrm{T}}$。
③另一种较简便的方法是构造辅助矩阵 $(\boldsymbol{A} \vdots \boldsymbol{I}_n)$：
（i）如果 $a_{11}\neq 0$，对辅助矩阵整体做行变换，可以将 $\boldsymbol{A}$ 的部分由上至下化为上三角阵，此时 $\boldsymbol{I}$ 的部分就是 $\boldsymbol{P}^{\mathrm{T}}$，$\boldsymbol{A}$ 的部分就是 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}$。
（ii）如果 $a_{11}=0$，可以通过初等变换使得 $a_{11}\neq 0$（$\boldsymbol{I}$ 的部分也一起变）；这一步做完之后，要对 $\boldsymbol{A}$ 的部分实施对应的列变换，但是这时 $\boldsymbol{I}$ 的部分不要变（相当于预先对 $\boldsymbol{A}$ 做一次合同变换；暂时略去对这一做法的解释）。当 $\boldsymbol{A}$ 的部分化为上三角阵时，$\boldsymbol{I}$ 的部分就是 $\boldsymbol{P}^{\mathrm{T}}$，但是 $\boldsymbol{A}$ 的部分不是 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}$（因为这个部分之前做了一次列变换）；要从头开始计算 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$，而不是默认 $\boldsymbol{A}$ 的部分为 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}$，直接右乘 $\boldsymbol{P}$。
7、正定二次型与正定矩阵
（1）惯性指数与惯性定理

①对于一个二次型的标准型，称其正系数的个数 p 为正惯性指数，负系数的个数 q 为负惯性指数，$p - q$ 称为符号差。
②惯性定理：同一个二次型的各个标准型有相同的正惯性指数和相同的负惯性指数（自然也有相同多的零系数）。
（2）正定二次型：若 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\geqslant 0$，当且仅当 $\boldsymbol{x}=\boldsymbol{0}$ 时取等号，则称 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为正定二次型，其相伴矩阵称为正定矩阵。
（3）若 $\boldsymbol{x}\neq\boldsymbol{0}$ 时 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\geqslant 0$ 也可取等号，则 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为半正定；
若 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\leqslant 0$，当且仅当 $\boldsymbol{x}=\boldsymbol{0}$ 时取等号，则 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为负定；
若 $\boldsymbol{x}\neq\boldsymbol{0}$ 时 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\leqslant 0$ 也可取等号，则 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为半负定。
（4）顺序主子式：矩阵的 k 阶顺序主子式就是其左上角的 k 阶子式。
（5）矩阵正定的充要条件（各个命题彼此等价）
①$\boldsymbol{A}$ 为正定矩阵；
②$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的正惯性指数为 n；
③$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的规范型为 $y_1^2+...+y_n^2$（系数全为+1）；
④$\boldsymbol{A}$ 的特征值均大于 0；
⑤$\boldsymbol{A}$ 的所有主子式均大于 0；
⑥$\boldsymbol{A}$ 的所有顺序主子式均大于 0；
⑦存在**唯一的**对角元素全正的下三角阵 $\boldsymbol{L}$ 使 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}$。（Cholesky 分解）
（6）若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 均为正定矩阵，则 $\boldsymbol{A}+\boldsymbol{B}$ 也正定；当实数 $c>0$ 时，$c\boldsymbol{A}$ 也正定；$\boldsymbol{AB}$ 不一定正定，这是因为 $\boldsymbol{AB}$ 不一定是对称阵，也就不一定能成为正定矩阵。（注：正定矩阵是相伴矩阵，因此 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为正定二次型不等价于 $\boldsymbol{A}$ 为正定矩阵，因为 $\boldsymbol{A}$ 可能并不是对称阵；这个看似矛盾的结论只是由于正定矩阵的定义比较严格而已——必须是对称阵/相伴矩阵。）
（7）若 $\boldsymbol{A}$ 为正定矩阵，则必存在唯一的正定矩阵 $\boldsymbol{B}$ 使 $\boldsymbol{A}=\boldsymbol{B}^2$。也记 $\boldsymbol{B}=\sqrt{\boldsymbol{A}}$。

*矩阵 $\boldsymbol{B}$ 的寻找：存在正交阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & ... & \\ & & & \lambda_n\end{pmatrix}$，因此可以构造

$\boldsymbol{B}=\boldsymbol{P}\begin{pmatrix}\sqrt{\lambda_1} & & \\ & ... & \\ & & \sqrt{\lambda_n}\end{pmatrix}\boldsymbol{P}^{\mathrm{T}}$。（可以保证 $\lambda_1,\lambda_2,...,\lambda_n>0$，因为它们就是正定矩阵 $\boldsymbol{A}$ 的特征值）
（8）同时相似与同时合同
①同时相似于实对角阵：存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 与 $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}$ 都是实对角阵。
②同时合同于实对角阵：存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 与 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{P}$ 都是实对角阵。
③方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 同时相似于实对角阵的充要条件是 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可交换（相易），即 $\boldsymbol{AB}=\boldsymbol{BA}$。
④若 $\boldsymbol{A}$ 正定，$\boldsymbol{B}$ 为同阶实对称阵，则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 同时合同于实对角阵。
（9）QR 分解：任意可逆实矩阵 $\boldsymbol{A}$ 均可用唯一的方式分解成 $\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}$，其中 $\boldsymbol{Q}$ 为正交阵，$\boldsymbol{R}$ 为对角元素全正的上三角阵。

附录

1、Cholesky 分解的步骤——归纳法

已知正定矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$。

（1）对其左上角元素 Cholesky 分解：$(a_{11})_{1\times1}=(\sqrt{a_{11}})(\sqrt{a_{11}})^{\mathrm{T}}$. 得到 $\boldsymbol{L}_1=(\sqrt{a_{11}})$.

（2）假设已经知道了 $\boldsymbol{A}$ 左上角的 k 阶主子阵（顺序主子阵）的 Cholesky 分解：

$\boldsymbol{A}_k=\boldsymbol{L}_k\boldsymbol{L}_k^{\mathrm{T}}$. 若下一个顺序主子阵 $\boldsymbol{A}_{k+1}=\begin{pmatrix}\boldsymbol{A}_k & \boldsymbol{\alpha}\\ \boldsymbol{\alpha}^{\mathrm{T}} & \mu\end{pmatrix}$，则可以构造 $\boldsymbol{L}_{k+1}=\begin{pmatrix}\boldsymbol{A}_k & \boldsymbol{0}\\ \boldsymbol{x}^{\mathrm{T}} & b\end{pmatrix}$, 使得

$\boldsymbol{A}_{k+1}=\boldsymbol{L}_{k+1}\boldsymbol{L}_{k+1}^{\mathrm{T}}$. 其中 $\begin{cases}\boldsymbol{x}=\boldsymbol{L}_k^{-1}\boldsymbol{\alpha}\\ b=\sqrt{\mu-\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_k^{-1}\boldsymbol{\alpha}}\end{cases}$。

（3）如果上述过程可以一直做到底，则完成了整个矩阵的 Cholesky 分解；如果一开始不知道矩阵 $\boldsymbol{A}$ 是否正定，而中间某个步骤做不下去（例如无法开平方，一些需要求逆的矩阵不可逆），那么说明 $\boldsymbol{A}$ 不是正定的。

2、QR 分解的步骤

①将矩阵 $\boldsymbol{A}$ 进行列分块：$\boldsymbol{A}=(\boldsymbol{A}_1,\dots,\boldsymbol{A}_n)$。

②将向量 $\boldsymbol{A}_1,\dots,\boldsymbol{A}_n$ 正交化：$\boldsymbol{B}_1=\boldsymbol{A}_1$，$i\geqslant 2$ 时 $\boldsymbol{B}_i=\boldsymbol{A}_i-\sum\limits_{j=1}^{i-1}\dfrac{(\boldsymbol{A}_i,\boldsymbol{B}_j)}{(\boldsymbol{B}_j,\boldsymbol{B}_j)}\boldsymbol{B}_j$

将 $\boldsymbol{B}_1,\dots,\boldsymbol{B}_n$ 单位化，得到一组标准正交基 $\boldsymbol{q}_1,\dots,\boldsymbol{q}_n$；它们排成的矩阵 $(\boldsymbol{q}_1,\dots,\boldsymbol{q}_n)$ 就是需要的正交矩阵 $\boldsymbol{Q}$。

③矩阵 $\boldsymbol{R}$ 在正交化与单位化的过程中得到：

$$\boldsymbol{q}_1=\frac{\boldsymbol{A}_1}{\|\boldsymbol{A}_1\|}\Rightarrow\boldsymbol{A}_1=\|\boldsymbol{A}_1\|\boldsymbol{q}_1\equiv r_{11}\boldsymbol{q}_1;$$

$$\boldsymbol{q}_2=\frac{1}{\|\boldsymbol{B}_2\|}\left[\boldsymbol{A}_2-\frac{(\boldsymbol{A}_2,\boldsymbol{B}_1)}{(\boldsymbol{B}_1,\boldsymbol{B}_1)}\boldsymbol{B}_1\right]=\frac{1}{\|\boldsymbol{B}_2\|}\left[\boldsymbol{A}_2-\frac{(\boldsymbol{A}_2,\boldsymbol{B}_1)}{(\boldsymbol{B}_1,\boldsymbol{B}_1)}\|\boldsymbol{A}_1\|\boldsymbol{q}_1\right]\Rightarrow\boldsymbol{A}_2=\frac{(\boldsymbol{A}_2,\boldsymbol{B}_1)}{(\boldsymbol{B}_1,\boldsymbol{B}_1)}\|\boldsymbol{A}_1\|\boldsymbol{q}_1+\|\boldsymbol{B}_2\|\boldsymbol{q}_2\equiv r_{12}\boldsymbol{q}_1+r_{22}\boldsymbol{q}_2.$$

以此类推，有 $(\boldsymbol{A}_1,\dots,\boldsymbol{A}_n)=(\boldsymbol{q}_1,\dots,\boldsymbol{q}_n)\begin{pmatrix}r_{11} & r_{12} & \cdots & r_{1n}\\ & r_{22} & \cdots & r_{2n}\\ & & \cdots & \\ & & & r_{nn}\end{pmatrix}$。

可以得到 $\boldsymbol{R}=\begin{pmatrix}r_{11} & r_{12} & \cdots & r_{1n}\\ & r_{22} & \cdots & r_{2n}\\ & & \cdots & \\ & & & r_{nn}\end{pmatrix}$。对角元素 $r_{ii}=\|\boldsymbol{B}_i\|>0$。（$1\leqslant i\leqslant n$）

[参考资料]
[1]复旦大学数学科学学院朱胜林老师《线性代数与概率统计》课程
[2]金路, 童裕孙, 於崇华, 张万国. 高等数学（第四版，上册）.高等教育出版社, 第 4-5 章
[3]金路. 线性代数同步辅导与复习提高. 复旦大学出版社.
[4]姚慕生, 吴泉水, 谢启鸿. 高等代数学（第三版）. 复旦大学出版社.