《线性代数》第一部分

南京大学数学学院 曹立元

2025 年 8 月 24 日创建，2025 年 10 月 11 日更新

1. 向量与矩阵的概念和基本运算

1.1. 向量与矩阵的概念

线形方程组

$$\begin{cases} 6x_1+3x_2+9x_3+9x_4=-1 \\ 8x_1-8x_2+9x_3+0x_4=8 \\ -7x_1-4x_2-7x_3+6x_4=6 \\ 8x_1+1x_2+9x_3-7x_4=9 \end{cases}$$

如果不想重复写 x、加号、等号那么多次，可以写成

$$\begin{bmatrix} 6 & 3 & 9 & 9 \\ 8 & -8 & 9 & 0 \\ -7 & -4 & -7 & 6 \\ 8 & 1 & 9 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -1 \\ 8 \\ 6 \\ 9 \end{bmatrix}.$$

令

$$\boldsymbol{A} = \begin{bmatrix} 6 & 3 & 9 & 9 \\ 8 & -8 & 9 & 0 \\ -7 & -4 & -7 & 6 \\ 8 & 1 & 9 & -7 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} -1 \\ 8 \\ 6 \\ 9 \end{bmatrix},$$

则上式可写为 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$.

这里 $\boldsymbol{A} \in \mathbb{R}^{4\times 4}$ 是 4×4 的实矩阵 (real matrix)，$\boldsymbol{x}, \boldsymbol{b} \in \mathbb{R}^{4\times 1}$ 是 4×1 的实矩阵，也是 4 维列向量。列向量 (column vector) 是只有一列的矩阵，行向量 (row vector) 是只有一行的矩阵，标量 (scalar) 是 1×1 的矩阵。

零向量 $\mathbf{0}$、全一向量 $\mathbf{1}$、沿第 i 个坐标轴的单位向量 $\boldsymbol{e}_i$

零矩阵 (zero matrix) $\mathbf{0}_{m\times n}$、方阵 (square matrix)、对角矩阵 (diagonal matrix)

1.2. 向量与矩阵的加、减、乘、幂

加、减、数乘都简单，关键是乘法不一样，还不满足交换律。对于向量

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^n \quad 和 \quad \boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$$

其点积 (dot product) $\boldsymbol{u}\cdot\boldsymbol{v}=u_1v_1+u_2v_2+\cdots+u_nv_n=\sum_{i=1}^n u_iv_i$. 这个计算有时被称也为内积，记为 $\langle\boldsymbol{u},\boldsymbol{v}\rangle$.

令 a,b 为任意实数，则

$$a\boldsymbol{u}+b\boldsymbol{v}=\begin{bmatrix}au_1\\au_2\\\vdots\\au_n\end{bmatrix}+\begin{bmatrix}bv_1\\bv_2\\\vdots\\bv_n\end{bmatrix}=\begin{bmatrix}au_1+bv_1\\au_2+bv_2\\\vdots\\au_n+bv_n\end{bmatrix}$$

是向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的线形组合 (linear combination)。

矩阵 $A\in\mathbb{R}^{m\times n}$ 与向量 $x\in\mathbb{R}^n$ 的相乘可被视为 A 中的行向量与 x 的点积，

$$Ax=\begin{bmatrix}(A\text{的第 1 行})\cdot x\\(A\text{的第 2 行})\cdot x\\\vdots\\(A\text{的第 }m\text{ 行})\cdot x\end{bmatrix}$$

也可被视为 A 中的列向量的线形组合

$$Ax=x_1(A\text{的第 1 列})+x_2(\boldsymbol{A}\text{的第 2 列})+\cdots+x_n(\boldsymbol{A}\text{的第 }n\text{ 列}).$$

单位矩阵(identity matrix) 是对角线元素全为 1 的对角矩阵，记为 $\boldsymbol{I}$，或加上下标 $\boldsymbol{I}_n$ 以表明其阶数。单位矩阵类似于实数乘法中的 1，满足 $\boldsymbol{IA}=\boldsymbol{A}$ 和 $\boldsymbol{AI}=\boldsymbol{A}$.

幂 (exponential)：$\boldsymbol{A}^k=\boldsymbol{A}^{k-1}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{k-1}$，$\boldsymbol{A}^0=\boldsymbol{I}$，只适用于方阵。

除法可以写成指数为负整数的幂，对于实数我们有 $a^{-1}a=1$，对于矩阵有 $\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{I}$，但这个比较复杂，我们以后讲。

1.3. 矩阵的转置

矩阵的转置 (transpose) 由上标 T 表达，是行列互换，即原本在第 i 行第 j 列上的元素转置后将出现在第 j 行第 i 列上。比如

$$\begin{bmatrix}2&4&6\\6&8&10\end{bmatrix}^T=\begin{bmatrix}2&6\\4&8\\6&10\end{bmatrix}.$$

转置具备以下性质：

1. $(\boldsymbol{A}^T)^T=\boldsymbol{A}$;
2. $(\boldsymbol{A}+\boldsymbol{B})^T=\boldsymbol{A}^T+\boldsymbol{B}^T$;
3. $(k\boldsymbol{A})^T=k\boldsymbol{A}^T\quad$ (k 是标量)；
4. $(\boldsymbol{AB})^T=\boldsymbol{B}^T\boldsymbol{A}^T$.

在欧里几得空间 (Euclidean space) 中，对于列向量 $\boldsymbol{u}, \boldsymbol{v}$，有 $\boldsymbol{u} \cdot \boldsymbol{v} = \langle \boldsymbol{u}, \boldsymbol{v} \rangle = \boldsymbol{u}^T \boldsymbol{v}$. 如果对为什么有这么多意义相同的表达方式感兴趣的话，可以看陶哲轩 (Terry Tao) 的这个帖子：https://mathoverflow.net/a/366118.

1.4. 《线形代数》的命名

这门课为什么叫线形代数?

首先代数是研究数、数量、关系、结构与代数方程（组）的通用解法及其性质的数学分支。比如中小学期间的代数包含了实数、四则运算、一元二次方程的求解等。这类初等代数的教学内容包含如 $y = f(x_1, x_2, x_3)$ 这种多元函数，但讲得不深，而线形代数这门课的一个重要意义就在于为我们今后研究、使用多元实向量函数 $f : \mathbb{R}^n \to \mathbb{R}^m$ 打下基础。

对于一元实函数 $f : \mathbb{R} \to \mathbb{R}$，它当且仅当对所有 $x, y, k \in \mathbb{R}$ 同时满足 $f(x + y) = f(x) + f(y)$ 和 $f(kx) = kf(x)$ 时被称为线形函数 (linear function)。对于多元实向量函数 $f : \mathbb{R}^n \to \mathbb{R}^m$，线形函数被定义为对所有 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ 和 $k \in \mathbb{R}$ 同时满足

$$f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y}) \quad \text{和} \quad f(k\boldsymbol{x}) = kf(\boldsymbol{x})$$

的函数，并且 f 是线形函数的充要条件是它可以用某个矩阵 $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ 表示为 $f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$。这就是为什么这门课学的是向量与矩阵，却叫线性代数。

2. 向量与矩阵及其基本运算的几何意义

2.1. 向量的几何意义

可以用来代表空间中一个点，每个元素代表一个坐标轴上的数值；也可以用来代表从原点指向那个点的向量；也可代表上述向量经过平移后的向量。

向量的长度：$\|\boldsymbol{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$.

向量的方向：向量本身带方向，但一般方向用单位向量（即长度为 1 的向量）来表示：$\dfrac{\boldsymbol{v}}{\|\boldsymbol{v}\|}$.

2.2. 向量运算的几何意义

向量相加：将向量 $\boldsymbol{v}$ 平移，将其本在原点的一端移至 $\boldsymbol{u}$ 点，本在 $\boldsymbol{v}$ 点的一端所处的位置则将是 $\boldsymbol{u}+\boldsymbol{v}$.

向量数乘：若 $k>0$，$k\boldsymbol{v}$ 是让向量方向不变，但长度变为原来的 k 倍；若 $k<0$，则是方向掉头，长度再变为原来的 $-k$ 倍。

$\boldsymbol{v}\cdot\boldsymbol{v} = \|\boldsymbol{v}\|^2$

向量的夹角：非零向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 之间的夹角 θ 满足 $\cos\theta = \dfrac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|\,\|\boldsymbol{v}\|}$ （余弦定理 law of cosines）.

垂直/正交向量：当 $\boldsymbol{u}\cdot\boldsymbol{v}=0$ 时，称向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 互相垂直 (perpendicular) 或正交 (orthogonal)。

投影长度：$\boldsymbol{u}\cdot\frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} = \|\boldsymbol{u}\|\cos\theta$.

Schwarz 不等式：$|\boldsymbol{u}\cdot\boldsymbol{v}| \le \|\boldsymbol{u}\|\,\|\boldsymbol{v}\|$.

三角不等式 (triangular inequality)：$\|\boldsymbol{u}+\boldsymbol{v}\| \le \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$，即三角形两边之和大于第三边。

2.3. 矩阵的几何意义

矩阵 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$ 可以被视为 m 个 $\mathbb{R}^n$ 空间中的向量（长度为 n 的向量），也可被视为 n 个 $\mathbb{R}^m$ 空间中的向量。

2.4. 线形空间

定义 1 (线形空间). 若一个向量的集合对其中任意两个元素 $\boldsymbol{u},\boldsymbol{v}$ 和任意实数 k 满足

$$(1)\ \ \boldsymbol{u}+\boldsymbol{v}\ \text{仍属于该集合，}\qquad (2)\ \ k\,\boldsymbol{v}\ \text{仍属于该集合，}$$

则称该集合为一个线性空间 (linear space) 或 向量空间 (vector space).

例：$\mathbb{R}^n$, 原点 $\{\boldsymbol{0}\}$, 穿过原点的任意直线，穿过原点的任意平面，...

若某线形空间的子集仍是线形空间，则称该子集为原空间的线形子空间 (linear subspace)，简称子空间。$\mathbb{R}^3$ 空间的子空间有：原点、穿过原点的任意直线、穿过原点的任意平面、和 $\mathbb{R}^3$ 本身。与子空间对应的概念是全空间，这是一个很模糊的概念，取决于应用场景，没有严谨的定义。一般来说，如果你的问题总共就 n 个变量，那么 $\mathbb{R}^n$ 就是这个问题的全空间。

一组向量所有线形组合的集合被称为这些向量的线形生成空间 (linear span). 比如对于向量集合

$\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m\} \subset \mathbb{R}^n$，其线形生成空间为

$$\operatorname{span}(\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m\}) = \left\{ \sum_{i=1}^{m} x_i \boldsymbol{a}_i : \ x_i \in \mathbb{R} \right\} \subseteq \mathbb{R}^n,$$

是 $\mathbb{R}^n$ 的子空间。

2.5. 矩阵乘法的几何意义 1、列空间、行空间

记 $\boldsymbol{a}_i$ 为矩阵 $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ 的第 i 列。$\boldsymbol{Ax}$ 可以看成 $\boldsymbol{A}$ 中列向量的线形组合

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n.$$

而这些列向量所有可能的组合所构成的集合被称为矩阵 $\boldsymbol{A}$ 的列空间 (column space):

$$\operatorname{span}(\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}) = \left\{ \sum_{i=1}^{n} x_i \boldsymbol{a}_i : \ x_i \in \mathbb{R} \right\} = \{\boldsymbol{Ax} : \boldsymbol{x} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m.$$

相应地，矩阵的行空间 (row space) 是

$$\left\{\boldsymbol{A}^T\boldsymbol{x} : \boldsymbol{x} \in \mathbb{R}^m\right\} \subseteq \mathbb{R}^n.$$

2.6. 线形函数的几何图像

从几何上看，线形函数 $f : \mathbb{R}^n \to \mathbb{R}^m$ 的图像属于 $\mathbb{R}^{n+m}$ 空间，是平直的，如直线、平面等，并且都过原点。

形如 $f(\boldsymbol{x}) = \boldsymbol{Ax} + \boldsymbol{b}$ 的函数被称为仿射函数 (affine function)，虽然对应的图像也是平直的，但未必过原点。然而，在很多线性代数的应用中，仿射函数往往就被称为线形函数，只有在需要强调 $\boldsymbol{b}$ 不一定是零向量时才可能使用仿射函数这一称呼。

2.7. 矩阵乘法的几何意义 2 与坐标变换

记 $\boldsymbol{a}_i$ 为矩阵 $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ 的第 i 行转置而成的列向量。$\boldsymbol{Ax}$ 可以看成 $\boldsymbol{x}$ 向 $\boldsymbol{A}$ 中各个行向量的投影

$$\begin{bmatrix} \boldsymbol{a}_1^T\boldsymbol{x} \\ \boldsymbol{a}_2^T\boldsymbol{x} \\ \vdots \\ \boldsymbol{a}_m^T\boldsymbol{x} \end{bmatrix}.$$

$\boldsymbol{A}$ 可看作对一个 $\mathbb{R}^n$ 空间的坐标变换。比如点 $\begin{pmatrix}1 & 1\end{pmatrix}^T$ 在由 $A = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$ 的行向量定义的新坐标系下的坐标是

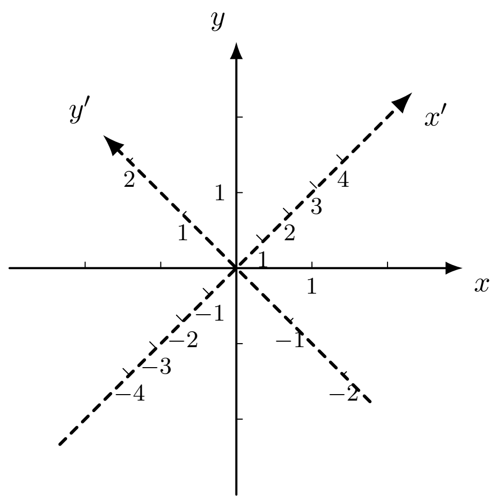
$$\begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix}; \qquad \text{(旋转)}$$

在由 $A=\begin{bmatrix}2&0\\0&1\end{bmatrix}$ 的行向量定义的新坐标系下的坐标是

$$\begin{bmatrix}2&0\\0&1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}2\\1\end{bmatrix}. \tag{拉伸}$$

先旋转再拉伸则是

$$\begin{bmatrix}2&0\\0&1\end{bmatrix}\begin{bmatrix}\sqrt{2}/2&\sqrt{2}/2\\-\sqrt{2}/2&\sqrt{2}/2\end{bmatrix}=\begin{bmatrix}\sqrt{2}&\sqrt{2}\\-\sqrt{2}/2&\sqrt{2}/2\end{bmatrix}.$$



注：这里两个矩阵相乘的顺序不能反过来，先拉伸再旋转所对应的矩阵不一样。因为 45 度的旋转是在原坐标系下定义的，拉伸后就不一样了。

坐标变换是线性的换元，比如上述矩阵可视为对实变量 x, y 做换元

$$\begin{aligned}x'&=\sqrt{2}x+\sqrt{2}y,\\ y'&=-\sqrt{2}x/2+\sqrt{2}y/2,\end{aligned}\qquad \text{即}\qquad \begin{bmatrix}x'\\y'\end{bmatrix}=\begin{bmatrix}\sqrt{2}&\sqrt{2}\\-\sqrt{2}/2&\sqrt{2}/2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}.$$

3. 关于矩阵的一些基本概念

3.1. 子阵

一个矩阵的子阵 (submatrix) 是原矩阵删掉几行几列后剩下的矩阵，比如下列矩阵在删掉第三行和第二、四列后剩下的矩阵就是原矩阵的一个子阵：

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \quad \to \quad \begin{bmatrix} 1 & 3 \\ 5 & 7 \end{bmatrix}.$$

主子阵 (principal submatrix)：留下的行和列的序号一致的子阵。如上述矩阵只留下 2、3 行和列的主子阵是

$$\begin{bmatrix} 6 & 7 \\ 10 & 11 \end{bmatrix}.$$

顺序主子阵 (leading principal submatrix)：只留前若干行和列的主子阵。这个概念不常见，有时人们会把它当作主子阵的定义。例：上述矩阵的顺序主子阵有

$$\begin{bmatrix} 1 \end{bmatrix}、\begin{bmatrix} 1 & 2 \\ 5 & 6 \end{bmatrix}、和 \begin{bmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{bmatrix}$$

3.2. 分块矩阵

$$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1s} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{A}_{r1} & \boldsymbol{A}_{r2} & \cdots & \boldsymbol{A}_{rs} \end{bmatrix} \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \cdots & \boldsymbol{B}_{1t} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} & \cdots & \boldsymbol{B}_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{B}_{s1} & \boldsymbol{B}_{s2} & \cdots & \boldsymbol{B}_{st} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^s \boldsymbol{A}_{1i}\boldsymbol{B}_{i1} & \sum_{i=1}^s \boldsymbol{A}_{1i}\boldsymbol{B}_{i2} & \cdots & \sum_{i=1}^s \boldsymbol{A}_{1i}\boldsymbol{B}_{it} \\ \sum_{i=1}^s \boldsymbol{A}_{2i}\boldsymbol{B}_{i1} & \sum_{i=1}^s \boldsymbol{A}_{2i}\boldsymbol{B}_{i2} & \cdots & \sum_{i=1}^s \boldsymbol{A}_{2i}\boldsymbol{B}_{it} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^s \boldsymbol{A}_{ri}\boldsymbol{B}_{i1} & \sum_{i=1}^s \boldsymbol{A}_{ri}\boldsymbol{B}_{i2} & \cdots & \sum_{i=1}^s \boldsymbol{A}_{ri}\boldsymbol{B}_{it} \end{bmatrix}$$

$$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \boldsymbol{A}_{13} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \boldsymbol{A}_{23} \end{bmatrix}^T = \begin{bmatrix} \boldsymbol{A}_{11}^T & \boldsymbol{A}_{21}^T \\ \boldsymbol{A}_{12}^T & \boldsymbol{A}_{22}^T \\ \boldsymbol{A}_{13}^T & \boldsymbol{A}_{23}^T \end{bmatrix}$$

3.3. 几类有特殊结构的矩阵

上三角矩阵 (upper triangular matrix)、下三角矩阵 (lower triangular matrix)、对称矩阵 (symmetric matrix) $\boldsymbol{A} = \boldsymbol{A}^T$、反对称矩阵 (skew-symmetric matrix) $\boldsymbol{A} = -\boldsymbol{A}^T$

行梯形矩阵 (row echelon form)、行简化梯形矩阵 (reduced row echelon form / row canonical form)、列梯形矩阵 (column echelon form)、列简化梯形矩阵 (reduced column echelon form / column canonical form)、

标准形矩阵 (canonical form) $\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}$

4. 初等变换

初等变换 (elementary operations) 分三种:

- 行(列)对调变换 (row/column switching): $r_i \leftrightarrow r_j (c_i \leftrightarrow c_j)$
- 行(列)数乘变换 (row/column multiplication): $kr_i(kc_i)$ 其中 $k \neq 0$
- 行(列)倍加变换 (row/column addition): $r_i + kr_j(c_i + kc_j)$

初等矩阵 (elementray matrices) 是对单位矩阵 I 进行一次初等变换后的矩阵(因此都是方阵),它们可以用来表示初等变换,包含初等对调矩阵、初等倍乘矩阵、初等倍加矩阵。

比如

$$\boldsymbol{P} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \begin{array}{l} \leftarrow \text{选择右边矩阵的第三行作为新的第一行} \\ \leftarrow \text{选择右边矩阵的第二行作为新的第二行} \\ \leftarrow \text{选择右边矩阵的第一行作为新的第三行} \end{array}$$

是对于有三行的矩阵的一三行对调变换 $r_1 \leftrightarrow r_3$ 所对应的行对调矩阵,$\boldsymbol{PA}$ 与 $\boldsymbol{A}$ 两个矩阵的关系是它们的一三行对调了;而

$$\boldsymbol{Q} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \begin{array}{l} \text{选择左边矩阵的第一列作为新的第一列} \\ \text{选择左边矩阵的第三列作为新的第二列} \\ \text{选择左边矩阵的第二列作为新的第三列} \end{array}$$

是对于有三列的矩阵的二三列对调变化 $c_2 \leftrightarrow c_3$,$\boldsymbol{AQ}$ 与 $\boldsymbol{A}$ 两个矩阵的关系是它们的一三行对调了。$r_i \leftrightarrow r_j$ 和 $c_i \leftrightarrow c_j$ 对应同一个矩阵,对调行还是列取决于从左侧还是右侧乘。

定理 1. 一个矩阵可以通过有限次行初等变化转化为行(简化)梯形矩阵;也可以通过有限次列初等变化转化为列(简化)梯形矩阵。(因此一个矩阵可以通过有限次初等变化转化为标准形。)

定理 2 (初等矩阵的逆矩阵). 对于任意一个初等变化(表示为初等矩阵 $\boldsymbol{Q}$),总存在另一个初等变换(表示为初等矩阵 $\boldsymbol{Q}^{-1}$)可将其逆转,使得 $\boldsymbol{Q}\boldsymbol{Q}^{-1} = \boldsymbol{I} = \boldsymbol{Q}^{-1}\boldsymbol{Q}$。

证明. 如果 $\boldsymbol{Q}$ 代表 $r_i \leftrightarrow r_j$ $(c_i \leftrightarrow c_j)$, 则 $\boldsymbol{Q}^{-1}$ 代表 $c_i \leftrightarrow c_j (c_i \leftrightarrow c_j)$;

$\boldsymbol{Q}$ 代表 kr_i (kc_i), 则 $\boldsymbol{Q}^{-1}$ 代表 $k^{-1}r_i$ $(k^{-1}c_i)$;

$\boldsymbol{Q}$ 代表 $r_i + kr_j$ $(c_j + kc_i)$, 则 $\boldsymbol{Q}^{-1}$ 代表 $r_i - kr_j$ $(c_j - kc_i)$. □

初等变换主要用于线性代数中的理论分析、手解线形方程组中,在这门课上很重要。然而在线性代数的应用中,人们一般会直接使用通过初等变换推导出的结论,初等变换基本不会出现。

5. 行列式

5.1. 行列式的定义

行列式 (determinant) 是一种由方阵向实数的映射 $\det:\mathbb{R}^{n\times n}\to\mathbb{R}$. 对于 1×1 的矩阵，其行列式就是其唯一元素的值。

余子式 (minor) M_{ij} 为一个矩阵去掉 i 行和 j 列后留下的子阵的行列式值；代数余子式 (cofactor) 为 $A_{ij}=(-1)^{i+j}M_{ij}$. 对于 $\boldsymbol{A}\in\mathbb{R}^{n\times n}$，行列式可通过以下迭代进行定义:

$$\sum_{i=1}^{n} a_{1i}A_{1i}.$$

对于二阶方阵 $\boldsymbol{A}$，其行列式为

$$\det(\boldsymbol{A})=\begin{vmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}.$$

若 $\boldsymbol{A}$ 为三阶方阵，其行列式为

$$\det(\boldsymbol{A})=\begin{vmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}=a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32} \\ -a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}.$$

此外还有一个基于逆序数 (signature of a permutation) 的 n 阶行列式的定义。

若 $\det(\boldsymbol{A})=0$，则称 $\boldsymbol{A}$ 是一个奇异矩阵 (singular matrix)；
若 $\det(\boldsymbol{A})\neq 0$，则称 $\boldsymbol{A}$ 是一个非奇异矩阵 (non-singular matrix)

5.2. 行列式的基本性质

行列式有些很重要的基本性质。它们的证明很繁琐，正常情况下不考。

定理 3. 对于 $\boldsymbol{A}\in\mathbb{R}^{n\times n}$，有 $\det(\boldsymbol{A})=\det(\boldsymbol{A}^T)$.

1. 三角矩阵的行列式的值为主对角线上元素的乘积。(书上例 1.2.2)
2. 奇数阶反对称矩阵的行列式的值为零。(书上例 1.2.5)
3. 对于分块矩阵 $\boldsymbol{M}=\begin{bmatrix} A & \mathbf{0}\\ C & B\end{bmatrix}$ 或 $\begin{bmatrix} A & C\\ \mathbf{0} & B\end{bmatrix}$，其中 $\boldsymbol{A},\boldsymbol{B}$ 为方阵（未必同阶），有 $\det(\boldsymbol{M})=\det(\boldsymbol{A})\det(\boldsymbol{B})$. （书上例 1.2.7)

定理 4. 对于 $\boldsymbol{A},\boldsymbol{B}\in\mathbb{R}^{n\times n}$，有 $\det(\boldsymbol{AB})=\det(\boldsymbol{A})\det(\boldsymbol{B})$. （书上例 2.2.7)

定理 5 (柯西-比内公式 (Cauchy–Binet formula)*). 对于 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$ 和 $\boldsymbol{B}\in\mathbb{R}^{n\times m}$，$\det(\boldsymbol{AB})=\sum_{S\in\binom{[n]}{m}}\det(\boldsymbol{A}_{[m],S})\det(\boldsymbol{B}_{S,[m]})$，其中 $[m]$ 代表有序数组 $1,2,\ldots,m$，$\binom{[n]}{m}$ 代表 $[n]$ 中所有长度为 m 的子数组所组成的集合，$\boldsymbol{A}_{[m],S}$ 代表 $\boldsymbol{A}$ 的 $[m]$ 行和 S 列所组成的矩阵。

5.3. 初等矩阵的行列式

根据基于逆序数的定义，不难得出初等矩阵有以下行列式值：

对调矩阵：-1；　　数乘矩阵：乘数 k；　　倍加矩阵：1.

结合定理 4，可得到

- 对调两行（列）的位置，行列式的值差一个负号（书上的定理 1.2.2）；
- 行列式的任一行（列）元素的公因子可以提到行列式外面（书上定理 1.2.5）；
- 将行列式的任一行（列）乘以 $k \in \mathbb{R}$ 再加到另一行（列）上去，行列式的值不变（书上定理 1.2.8）。

综上，初等变换不会改变矩阵的奇异性。

5.4. 行列式的几何意义

记 $\boldsymbol{a}_i$ 为矩阵 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 的第 i 行（列），则

$$\det(\boldsymbol{A}) \text{ 为平行多面体} \left\{ \sum_{i=1}^{n} \eta_i \boldsymbol{a}_i：\ \eta_i \in [0,1]\ \forall i \right\} \text{ 的体积乘以 } \pm 1。$$

对于二阶方阵，若从第一个向量沿着逆时针方向转到第二个向量的弧度小于 π，则乘 1；否则乘 -1.

对于任意小的 $\varepsilon > 0$，$\varepsilon \boldsymbol{I}$ 代表一个体积为 ϵ 的小正方体，而 $\boldsymbol{A}(\varepsilon \boldsymbol{I}) = \varepsilon \det(\boldsymbol{A})$. 故根据微积分中的划分区域的思想，用 $\boldsymbol{A}$ 乘以任意 $\boldsymbol{B}$，会使 $\boldsymbol{B}$ 所代表的平行多面体的体积变化 $\det(\boldsymbol{A})$ 倍，即 $\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$.

6. 线形方程组的求解

线形方程组 (system of linear equations, or 线性系统 linear system) $\boldsymbol{Ax}=\boldsymbol{b}$
其中 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$ 和 $\boldsymbol{b}\in\mathbb{R}^m$ 已知，$\boldsymbol{x}\in\mathbb{R}^n$ 未知。

6.1. 克莱姆法则

若 $\boldsymbol{A}\in\mathbb{R}^{n\times n}$，$\boldsymbol{a}_i$ 为其第 i 列，线形方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 可写为

$$\boldsymbol{A}\begin{bmatrix}1 & & & x_1 & & & \\ & 1 & & x_2 & & & \\ & & \ddots & \vdots & & & \\ & & & x_i & & & \\ & & & \vdots & \ddots & & \\ & & & x_{n-1} & & 1 & \\ & & & x_n & & & 1\end{bmatrix}=\begin{bmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{b} & \cdots & \boldsymbol{a}_n\end{bmatrix}$$

等式左边 $\boldsymbol{A}$ 乘以的矩阵是第 i 列被替换成 $\boldsymbol{x}$ 后的 $\boldsymbol{I}$，其行列式的值为 x_i；等式右边的矩阵是第 i 列被替换成 $\boldsymbol{b}$ 后的 $\boldsymbol{A}$，记为 $\boldsymbol{A}_i$. 根据定理 4，有 $\det(\boldsymbol{A})\,x_i=\det(\boldsymbol{A}_i)$. 因此，若 $\det(\boldsymbol{A})\neq 0$，则线形方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解可以写为

$$x_i=\frac{\det(\boldsymbol{A}_i)}{\det(\boldsymbol{A})}.$$

该等式被称为克莱姆法则 (Cramer's rule)，虽然可用来求解线形方程组，但是当 n 大时效率低下，其价值更多地体现在理论分析中。

6.2. 高斯消元法

高斯消元法 (Gaussian elimination) 对线形方程组的行做初等变换，以求将矩阵化为行简化梯形矩阵。

定义 2. 具有相同解集的两个方程组称为同解方程组。

定理 6 (行初等变换不改变线性方程组的解集). 对于线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 和初等矩阵 $\boldsymbol{P}$，则 $\{\boldsymbol{x}:\ \boldsymbol{Ax}=\boldsymbol{b}\}=\{\boldsymbol{x}:\ \boldsymbol{PAx}=\boldsymbol{Pb}\}$.

证明. 若 $\boldsymbol{x}$ 满足 $\boldsymbol{Ax}=\boldsymbol{b}$，则 $\boldsymbol{x}$ 显然也满足 $\boldsymbol{PAx}=\boldsymbol{Pb}$，故 $\{\boldsymbol{x}:\ \boldsymbol{Ax}=\boldsymbol{b}\}\subseteq\{\boldsymbol{x}:\ \boldsymbol{PAx}=\boldsymbol{Pb}\}$. 类似地，若 $\boldsymbol{x}$ 满足 $\boldsymbol{PAx}=\boldsymbol{Pb}$ 的一个解，则 $\boldsymbol{x}$ 也满足 $\boldsymbol{Ax}=\boldsymbol{P}^{-1}\boldsymbol{PAx}=\boldsymbol{P}^{-1}\boldsymbol{Pb}=\boldsymbol{b}$，故 $\{\boldsymbol{x}:\ \boldsymbol{Ax}=\boldsymbol{b}\}\supseteq\{\boldsymbol{x}:\ \boldsymbol{PAx}=\boldsymbol{Pb}\}$. □

上述定理证明了高斯消元法不会改变解集。

若多个线形方程组的矩阵一样，可以同时进行:

$$\begin{bmatrix}2 & 1 & 3\\ 3 & 1 & 2\\ 4 & 3 & 8\end{bmatrix}\begin{bmatrix}x_1 & y_1\\ x_2 & y_2\\ x_3 & y_3\end{bmatrix}=\begin{bmatrix}-1 & 4\\ 8 & -6\\ 6 & -3\end{bmatrix}.$$

6.3. LU 分解

当 $\boldsymbol{A}$ 不变，需要针对多个不同的 $\boldsymbol{b}$ 求解线形方程组，而且不能同时获得所有 $\boldsymbol{b}$ 的值时，为了避免重复计算，可以使用 LU 分解 (LU (lower–upper) decomposition) 将 $\boldsymbol{A}$ 分解为一个对角线元素全为 1 的下三角矩阵和一个上三角矩阵的积。这其实是通过列初等变换将矩阵 A 化为下三角矩阵，同时用上三角矩阵记录这些变换的逆过程。

令 $\boldsymbol{Q}_1, \boldsymbol{Q}_2, \ldots, \boldsymbol{Q}_k$ 为 k 个初等矩阵，则 LU 分解法：

$$\begin{aligned} \boldsymbol{Ax} &= \boldsymbol{b} \\ \underbrace{\boldsymbol{A}\boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_k}_{\boldsymbol{L}}\underbrace{\boldsymbol{Q}_k^{-1}\boldsymbol{Q}_{k-1}^{-1}\cdots\boldsymbol{Q}_1^{-1}}_{\boldsymbol{U}}\boldsymbol{x} &= \boldsymbol{b} \\ \boldsymbol{L}(\boldsymbol{Ux}) &= \boldsymbol{b} \\ \boldsymbol{Ly} = \boldsymbol{b} \qquad \boldsymbol{Ux} &= \boldsymbol{y} \end{aligned}$$

6.4. 解的数量

线性方程组可能没有解，可能有唯一解，也可能有无穷多个解，哪怕其变量数与等式数相等：

$$\begin{bmatrix} 1 & -2 & 0 \\ 1 & 1 & 3 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \tag{无解}$$

$$\begin{bmatrix} 1 & -2 & 0 \\ 1 & 1 & 3 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1/3 \end{bmatrix} \tag{无穷多个解}$$

6.5. 应用：线形插值

两点确认一条线，三点确认一个面。针对一个未知、但可以取值的多元函数 $f : \mathbb{R}^n \to \mathbb{R}$，利用 $n+1$ 个点 $\{\boldsymbol{x}_i\}_{i=1}^{n+1}$ 的函数值，可以通过求解线形方程组

$$\begin{bmatrix} 1 & \boldsymbol{x}_1^T \\ 1 & \boldsymbol{x}_2^T \\ \vdots & \vdots \\ 1 & \boldsymbol{x}_n^T \end{bmatrix} \begin{bmatrix} c \\ \boldsymbol{g} \end{bmatrix} = \begin{bmatrix} f(\boldsymbol{x}_1) \\ f(\boldsymbol{x}_2) \\ \vdots \\ f(\boldsymbol{x}_n) \end{bmatrix}$$

确认关于 f 的一个局部线形拟合模型 $\hat{f}(\boldsymbol{x}) = c + \boldsymbol{g} \cdot \boldsymbol{x}$. 这种拟合方法被称为线性插值 (linear interpolation)。