第一章　线性代数

1.1　线性映射和矩阵

1.1.1　线性映射

定义　<u>线性运算</u>指向量的加法与数乘。

定义　<u>向量空间</u>为带有线性运算的集合 $\mathbb{R}^m$ 被称为向量空间。

定义　映射 $f : \mathbb{R}^n \to \mathbb{R}^m$ 若满足

1. 任意 $\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^n$，都有 $f(\boldsymbol{x} + \boldsymbol{x'}) = f(\boldsymbol{x}) + f(\boldsymbol{x'})$
2. 任意 $\boldsymbol{x} \in \mathbb{R}^n, k \in \mathbb{R}$, 都有 $f(k\boldsymbol{x}) = kf(\boldsymbol{x})$

则 f 为从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的 <u>线性映射</u>。

定义　从 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的线性映射称为<u>线性变换</u>.

定义　若线性映射 f 有 $f(\boldsymbol{e_i}) = \boldsymbol{a_i}, \boldsymbol{e_i} \in \mathbb{R}^n, \boldsymbol{a_i} \in \mathbb{R}^m$，则矩阵 $A = [\boldsymbol{a_1}, \ldots, \boldsymbol{a_n}]$ 即为标准坐标向量下的 <u>线性映射的表示矩阵</u>，且满足 $A\boldsymbol{e_i} = \boldsymbol{a_i}$。

定理　（线性映射的线性运算）若矩阵 $\boldsymbol{A}, \boldsymbol{B}$ 表示 $\mathbb{R}^n \to \mathbb{R}^m$ 的线性映射，则 $\boldsymbol{A} + \boldsymbol{B}$ 与 $k\boldsymbol{A}$ 也是同样范畴上的线性映射。

定理　$\boldsymbol{AB}$ 表示线性映射 $\boldsymbol{A}, \boldsymbol{B}$ 的复合 $\boldsymbol{A} \circ \boldsymbol{B}$。需要注意，$\boldsymbol{AB} = 0$ 也不能推出 $\boldsymbol{A} = 0$ 或 $\boldsymbol{B} = 0$。

理解　矩阵乘法 $A\boldsymbol{x}$ 视为这样一种运算：按照 $\boldsymbol{x}$ 指定的系数，将 $\boldsymbol{A}$ 的列进行线性组合。

定义　<u>反对称矩阵</u>：$A = -A^{\mathrm{T}}$，反对称矩阵对角线必定为 0。

定义　若一个上三角矩阵对角线上全为 0，则称为<u>**严格上三角矩阵**</u>。

定义　将阶梯形矩阵，从下向上消元，并单位化<u>**主元**</u>，得到的矩阵每个非零行上，主变量为 1 而其他列的元素均为 0，称这个矩阵为 <u>**行简化阶梯形矩阵**</u>。

定理　对于 $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^m$，$\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w} = trace(\boldsymbol{w}\boldsymbol{v}^{\mathrm{T}})$。

1.1.2 线性方程式组

定理　对于方程组 $A\boldsymbol{x} = \boldsymbol{b}$，将 $[A\ \boldsymbol{b}]$ 简化为阶梯型后，若

1. $[A\ \boldsymbol{b}]$ 阶梯数比 A 多 $\mathbf{1}$，则方程组无解（$0 \neq 1$）
2. $[A\ \boldsymbol{b}]$ 阶梯数与 A 相等，则方程组有解
   (a) 若阶梯数等于未知数个数，则有唯一解
   (b) 若阶梯数小于未知数个数，则有无穷多组解

定义　$A\boldsymbol{x} = \boldsymbol{b}$ 称为<u>**齐次线性方程组**</u>，$\vec{0}$ 为其<u>**平凡解**</u>，除此之外的解称为<u>**非平凡解**</u>。

1.1.3 可逆矩阵

定义　设 A 为 n 阶方阵，若存在 n 阶方阵 B，使得 $AB = BA = I_n$，则 A 为 <u>**可逆矩阵**</u>或<u>**非奇异矩阵**</u>，B 为 A 的逆。

定理　以下命题等价

1. A 可逆
2. 任意 $\boldsymbol{b} \in \mathbb{R}^n$，$A\boldsymbol{x} = \boldsymbol{b}$ 的解唯一
3. 其次方程组 $A\boldsymbol{x} = 0$ 仅有零解
4. A 对应的阶梯形矩阵有 n 个主元
5. A 对应的行简化阶梯形矩阵是 I_n
6. A 能表示为有限个初等矩阵的乘积（即消元至行简化阶梯形矩阵的逆过程）

定义　若矩阵 $A = [a_{ij}]_{n\times n}$ 对于 $i = 1, 2, \ldots, n$ 都有 $|a_{ii}| > \sum_{j\neq i} |a_{ij}|$，则称其为 <u>**（行）对角占优矩阵**</u>。

定理　对角占优矩阵必然可逆。

定义　若矩阵 A 通过若干初等**行变换**可以变为矩阵 B，则称 A, B <u>**左相抵**</u>。即存在可逆矩阵 P，使得 $PA = B, A = P^{-1}B$。所有和 A 相抵的矩阵中，最简单的是其行简化阶梯形，它被称

为 A 的左相抵标准形。

定理　　左相抵构成等价关系。

定理　　**(Sherman-Morrison)** 设 A 为 n 阶可逆方阵，$\boldsymbol{u},\boldsymbol{v}$ 为 n 阶向量，则 $A+\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ 可逆 $\iff 1+\boldsymbol{v}^{\mathrm{T}}A^{-1}\boldsymbol{u}\neq 0$，且此时

$$(A+\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}})^{-1}=A^{-1}-\frac{A^{-1}\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}A^{-1}}{1+\boldsymbol{v}^{\mathrm{T}}A^{-1}\boldsymbol{u}} \tag{1.1}$$

若将 $\boldsymbol{u},\boldsymbol{v}$ 改为 $n\times k$ 的矩阵，则类似地有

$$(A+\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}})^{-1}=A^{-1}-A^{-1}\boldsymbol{u}(I_k+\boldsymbol{v}^{\mathrm{T}}A^{-1}\boldsymbol{u})^{-1}\boldsymbol{v}^{\mathrm{T}}A^{-1} \tag{1.2}$$

1.1.4　LU 分解

定理　　若 n 阶方阵 A 仅通过倍加矩阵做行变化即可化为阶梯形，则存在单位下三角矩阵（主对角线均为 1）L 与上三角矩阵 U，使得 $A=LU$，此即 **LU 分解**。

定义　　方阵 A 左上角的 $k\times k$ 块为第 k 个顺序主子阵。

定理　　可逆矩阵 A 存在 LU 分解，当且仅当 A 的所有顺序主子阵均可逆，此时 LU 分解唯一。（满足在消元过程中不需要行的调换）

定理　　若可逆矩阵 A 存在 LU 分解，则存在对角线均不为 0 的对角矩阵 D、单位下三角矩阵 L、单位上三角矩阵 U，满足 $A=LDU$，且该分解**唯一**。此即 **LDU 分解**。

定理　　若可逆对称阵 A 有 LDU 分解，则 $L=U^{\mathrm{T}}$。

定理　　可逆矩阵 A 存在分解 $A=PLU$，P 为置换矩阵，显然该分解不唯一。

技巧　　将 A 分解成对称阵 $X=\frac{1}{2}(A+A^{\mathrm{T}})$ 与反对称阵 $Y=\frac{1}{2}(A-A^{\mathrm{T}})$。

1.2　子空间和维数

1.2.1　基本概念

定义　　映射 $A:\boldsymbol{x}\mapsto A\boldsymbol{x}\in\mathbb{R}^{m\times n}$ 的像集

$$\mathcal{R}(A)=\{A\boldsymbol{x}|\boldsymbol{x}\in\mathbb{R}^n\}\subseteq\mathbb{R}^m \tag{1.3}$$

显然 $\mathcal{R}(A)$ 必是子空间，即 A 的<u>列（向量）空间</u>。

命题　$\mathcal{R}(A)=\mathbb{R}^m \iff A$ 为满射。

定义　映射 $A:\boldsymbol{x}\mapsto A\boldsymbol{x}\in\mathbb{R}^{m\times n}$ 的<u>原像</u>

$$\mathcal{N}(A)=\{\boldsymbol{x}\in\mathbb{R}^n|A\boldsymbol{x}=\mathbf{0}\}\in\mathbb{R}^n \tag{1.4}$$

显然 $\mathcal{N}(A)$ 必是子空间，即 A 的<u>零空间（也称解空间）</u>。

命题　$\mathcal{N}(A)=\{\mathbf{0}\} \iff A$ 是单射 $\iff \forall\boldsymbol{x}_1\neq\boldsymbol{x}_2,\ A(\boldsymbol{x}_1-\boldsymbol{x}_2)\neq 0$。

定义　将 $\boldsymbol{a}_1,\dots,\boldsymbol{a}_n$ 的全体线性组合为 $\mathbb{R}^n$ 的子空间，记为

$$\operatorname{span}(\boldsymbol{a}_1,\dots,\boldsymbol{a}_n):=\{k_1\boldsymbol{a}_1,\dots,k_n\boldsymbol{a}_n|k_1,\dots,k_n\in\mathbb{R}\} \tag{1.5}$$

定理　$A=\begin{bmatrix}\boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n\end{bmatrix}\in\mathbb{R}^{n\times n}$ 可逆 $\iff \boldsymbol{a}_1,\dots,\boldsymbol{a}_n$ 是 $\mathbb{R}^n$ 的一组基。

1.2.2　基和维数

定义　向量组 S 的任一<u>极大线性无关组</u>中向量的个数称为 S 的<u>秩 (rank)</u>；子空间 $\mathcal{M}$ 的一组基的向量数目称为<u>维数</u>，记为 $\dim\mathcal{M}$。

定理　<u>（基存在定理）</u>给定 $\mathbb{R}^m$ 的子空间 $\mathcal{M}\neq\{0\}$，则 $\mathcal{M}$ 存在一组基，且基向量个数不超过 m。

定理　<u>（基扩充定理）</u>若 $\mathcal{M}\subseteq\mathcal{N}$，则 $\mathcal{M}$ 的任意一组基都能够<u>扩充</u>成 $\mathcal{N}$ 的一组基。

定理　对于方阵 $A\in\mathbb{R}^{n\times n}$，$A$ 可逆 $\iff A$ 是单射 $\iff A$ 是满射。

1.2.3　矩阵的秩

定义　执行 Gauss 消去法将矩阵化为阶梯型后，主元所在的列对应主变量的系数，称为<u>主列</u>；其他列对应自由变量的系数，称为<u>自由列</u>。

定理　若矩阵 $A=\begin{bmatrix}\boldsymbol{a}_1 & \dots & \boldsymbol{a}_n\end{bmatrix}$ 与 $B=\begin{bmatrix}\boldsymbol{b}_1 & \dots & \boldsymbol{b}_n\end{bmatrix}$ 左相抵，则

1. 部分组 $\boldsymbol{a}_{i_1},\dots,\boldsymbol{a}_{i_r}$ 线性无关，当且仅当 $\boldsymbol{b}_{i_1},\dots,\boldsymbol{b}_{i_r}$ 线性无关。
2. $\boldsymbol{a}_j=k_1\boldsymbol{a}_{i_1}+\cdots+k_r\boldsymbol{a}_{i_r} \iff \boldsymbol{b}_j=k_1\boldsymbol{b}_{i_1}+\cdots+k_r\boldsymbol{b}_{i_r}$

二级结论

- $\operatorname{rank}(A)+\operatorname{rank}(B)=\operatorname{rank}\left(\begin{bmatrix}A & O\\ O & B\end{bmatrix}\right)\le\operatorname{rank}\left(\begin{bmatrix}A & X\\ O & B\end{bmatrix}\right)$。因而当 A,B 可逆时，上述两个分块矩阵都可逆
- $\operatorname{rank}(A+B)\le\operatorname{rank}(A)+\operatorname{rank}(B)$
- $\operatorname{rank}(AB)\le\min\{\operatorname{rank}(A),\operatorname{rank}(B)\}$
- 反对称矩阵的秩必定是偶数
- $\mathcal{R}(AB)\subseteq\mathcal{R}(A)$

定义　$\mathcal{R}(A^{\mathrm{T}})$ 为 A 的行（向量）空间；$\operatorname{rank}(A)=\operatorname{rank}(A^{\mathrm{T}})$；$\operatorname{rank}(A)=n$ 定义为列满秩，同理有行满秩。

定理　对于 $A\in\mathbb{R}^{m\times n}$，

- A 是满射 $\iff$ A 行满秩
- A 是单射 $\iff$ A 列满秩

定义　通过行列初等变换，能将矩阵 A 变换为相抵标准型 $\begin{bmatrix}I_r & O\\ O & O\end{bmatrix}$。

1.2.4　线性方程组的解

方法　（求零空间的一组基）矩阵 $A\in\mathbb{R}^{m\times n}$，求其行简化阶梯形，考虑依次将各个自由元取为 1，同时维持其他自由元为 0（即只让其中一个自由列在线性组合中系数非零），由于主列都是单位向量，从而容易根据自由列的系数计算出零解。这样求出的解（由于自由元的取值特点）构成零空间的一组基，继而有 $\dim\mathcal{N}(A)=n-\operatorname{rank}(A)$。

1.3　内积和正交性

1.3.1　基本概念

定义　$\boldsymbol{b}$ 向 $\boldsymbol{a}$ 的垂直投影为 $\hat{x}=\frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}$；向量 $\frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}\boldsymbol{a}$ 称为向量 $\boldsymbol{b}$ 向直线 $\operatorname{span}(\boldsymbol{a})$ 的投影。

定理　**（Cauchy-Schwarz 不等式）** $|\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}|\le\|a\|\|b\|$。

定义　设 $\mathcal{M}$ 是 $\mathbb{R}^n$ 的子空间，若它的一组基是正交向量组，则称之为 $\mathcal{M}$ 的一组正交基；若是正交单位向量组，则称为 $\mathcal{M}$ 的标准正交基。

定义 <u>**（Gram-Schmidt 正交化）**</u>从 $\mathcal{M}$ 的任意一组基 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$ 出发，执行如下操作

$$\widetilde{\boldsymbol{q}}_k = \boldsymbol{a}_k - \sum_{j=1}^{k-1} \frac{\widetilde{\boldsymbol{q}}_j^{\mathrm{T}} \boldsymbol{a}_k}{\widetilde{\boldsymbol{q}}_j^{\mathrm{T}} \widetilde{\boldsymbol{q}}_j} \widetilde{\boldsymbol{q}}_j \quad (k = 1, 2, \ldots, r) \tag{1.6}$$

即得到一组正交基，最后再通过单位化，得到一组标准正交基。

1.3.2　正交矩阵和 QR 分解

定义 若方阵 Q 满足 $Q^{\mathrm{T}}Q = I_n$，则称 Q 为 n 阶<u>**正交矩阵**</u>；Q 的行、列向量各自构成 $\mathbb{R}^n$ 的一组标准正交基；多个正交矩阵的积也为正交矩阵。

定理 方阵 Q 的以下叙述等价

1. Q 是正交矩阵
2. Q 为<u>**保距变换**</u>，即 $\|Q\boldsymbol{x}\| = \|\boldsymbol{x}\|$
3. Q 为<u>**保内积变换**</u>，即 $Q\boldsymbol{x} \cdot Q\boldsymbol{y} = \boldsymbol{x} \cdot \boldsymbol{y}$

保距变换一定也是<u>**保角变换**</u>。

定义 <u>**（Givens 变换）**</u>在 $\boldsymbol{e}_i - \boldsymbol{e}_j$ 平面上转角 θ 的旋转变换的矩阵为

$$\begin{bmatrix} \ddots & & & & & & \\ & 1 & & & & & \\ & & \cos\theta & & -\sin\theta & & \\ & & & 1 & & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & \sin\theta & & \cos\theta & & \\ & & & & & 1 & \\ & & & & & & \ddots \end{bmatrix} \tag{1.7}$$

定义 <u>**（Householder 变换）**</u>关于 $\mathbb{R}^n$ 中的单位法向量所确定的超平面 $\mathcal{N}(\boldsymbol{v}^{\mathrm{T}})$ 进行反射变换的矩阵为

$$H_{\boldsymbol{v}} = I_n - 2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}} \tag{1.8}$$

不难注意到 $\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}\boldsymbol{w}$ 是 $\boldsymbol{w}$ 向 $\mathrm{span}(\boldsymbol{v})$ 的投影，因此 $H_{\boldsymbol{v}}$ 的效果是将 $\boldsymbol{w}$ 中与 $\boldsymbol{v}$ 共线的成分反向。

定义 <u>**（QR 分解）**</u>设 $A = \begin{bmatrix} \boldsymbol{a}_1 & \ldots & \boldsymbol{a}_n \end{bmatrix}$ 为 n 可逆矩阵。在对 $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ 进行 Gram-Schmidt 正交化过程的中，本质上是进行多次的列初等变换，故可以借此对 A 进行表达

$$\boldsymbol{a}_k = \widetilde{\boldsymbol{q}}_k + \sum_{j=1}^{k-1} \frac{\widetilde{\boldsymbol{q}}_j^{\mathrm{T}} \boldsymbol{a}_k}{\widetilde{\boldsymbol{q}}_j^{\mathrm{T}} \widetilde{\boldsymbol{q}}_j} \widetilde{\boldsymbol{q}}_j \quad (k = 1, 2, \ldots, r) \tag{1.9}$$

$$A = \widetilde{Q}\widetilde{R} = \begin{bmatrix}\widetilde{\boldsymbol{q}}_1 & \dots & \widetilde{\boldsymbol{q}}_n\end{bmatrix}\begin{bmatrix}1 & \frac{\widetilde{\boldsymbol{q}}_1^{\mathrm{T}}\boldsymbol{a}_2}{\widetilde{\boldsymbol{q}}_1^{\mathrm{T}}\widetilde{\boldsymbol{q}}_1} & \cdots & \frac{\widetilde{\boldsymbol{q}}_1^{\mathrm{T}}\boldsymbol{a}_n}{\widetilde{\boldsymbol{q}}_1^{\mathrm{T}}\widetilde{\boldsymbol{q}}_1} \\ & 1 & \ddots & \vdots \\ & & \ddots & \frac{\widetilde{\boldsymbol{q}}_{n-1}^{\mathrm{T}}\boldsymbol{a}_n}{\widetilde{\boldsymbol{q}}_{n-1}^{\mathrm{T}}\widetilde{\boldsymbol{q}}_{n-1}} \\ & & & 1\end{bmatrix} \tag{1.10}$$

可以进一步将 Q 单位化，得到

$$A = Q\ \mathrm{diag}(\|\widetilde{\boldsymbol{q}}_i\|)\widetilde{R} = QR \tag{1.11}$$

定理　设 A 为 n 阶可逆矩阵，则存在唯一的分解 $A = QR$，其中 Q 为正交矩阵，R 为对角元均为正数的上三角矩阵。

定义　若矩阵 Q 满足 $Q^{\mathrm{T}}Q = I_n$，则称之为 **列正交矩阵**。

定理　对于 $m \times n$ 的矩阵 A，其中 $m \geq n$，则

1. **（简化 QR 分解）**存在 $m \times n$ 列正交矩阵 Q_1 和具有非负对角元的 n 阶上三角矩阵 R_1，使得 $A = Q_1R_1$
2. **（QR 分解）**进一步的，存在 m 阶正交矩阵 Q 和 $m \times n$ 矩阵 R，使得 $A = QR = \begin{bmatrix}Q_1 & Q_2\end{bmatrix}\begin{bmatrix}R_1 \\ O\end{bmatrix}$

1.3.3　正交投影

定义　**（子空间正交）**若子空间 $\mathcal{M}$ 的任意向量与 $\mathcal{N}$ 中的任意向量都正交，则称为 $\mathcal{M}$ 与 $\mathcal{N}$ 正交，记为 $\mathcal{M} \perp \mathcal{N}$。

若 $\mathcal{M}, \mathcal{N} \in \mathbb{R}^n, \mathcal{M} \perp \mathcal{N}, \mathcal{M} \cup \mathcal{N} = \mathbb{R}^n$，则称 $\mathcal{N}$ 为 $\mathcal{M}$ 的**正交补**，记为 $\mathcal{N} = \mathcal{M}^{\perp}$。

定理　对于 $A \in \mathbb{R}^{m\times n}$，有

1. $\mathcal{R}(A^{\mathrm{T}})^{\perp} = \mathcal{N}(A), \mathcal{R}(A)^{\perp} = \mathcal{N}(A^{\mathrm{T}})$（矩阵导出的四个子空间的关系）
2. $\mathcal{R}(A^{\mathrm{T}}A) = \mathcal{R}(A^{\mathrm{T}}), \mathcal{N}(A^{\mathrm{T}}A) = \mathcal{N}(A)$

TBD　正交投影

1.4　行列式

1.4.1　行列式函数

定义　定义在全体 n 阶方阵上的函数 δ，如果满足如下性质：

1. 列多线性性：$\delta(\cdots, k\boldsymbol{a}_i + k'\boldsymbol{a}_i', \cdots) = k\delta(\cdots, \boldsymbol{a}_i, \cdots) + k'\delta(\cdots, \boldsymbol{a}_i', \cdots)$
2. 列反对称性：交换任意两列，函数符号反转
3. 单位化条件：$\delta(I_n) = 1$

则称 δ 为一个 n 阶**行列式函数**，该函数存在且唯一。容易证明 $\det(AB) = \det(A)\det(B)$。

定义 **Vandermonde 矩阵**及其行列式

$$\begin{vmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \cdots & \lambda_n^{n-1} \end{vmatrix} = \prod_{1 \le j < i \le n} (\lambda_i - \lambda_j) \tag{1.12}$$

1.4.2 行列式的展开式

定义 给定 n 阶方阵 A，令 $A\binom{i}{j}$ 表示从 A 中划去第 i 行第 j 列后得到的 $n-1$ 阶方阵，称 $M_{ij} = \det\left(A\binom{i}{j}\right)$ 为元素 a_{ij} 的**余子式**；而 $C_{ij} = (-1)^{i+j}M_{ij}$ 为元素 a_{ij} 的**代数余子式**。

定理 行列式按第一列展开 $\det(A) = a_{11}C_{11} + \cdots + a_{n1}C_{n1}$

定理 令 A 的第 i 列为 $\boldsymbol{a}_i$，记第 j 列元素的代数余子式组成的向量 $\boldsymbol{c}_j = \begin{bmatrix} C_{1j} \\ \vdots \\ C_{nj} \end{bmatrix}$，则

$$\boldsymbol{a}_{j'}^{\mathrm{T}}\boldsymbol{c}_j = \begin{cases} 0, & j \ne j' \\ \det(A), & j = j' \end{cases} \tag{1.13}$$

定义 对矩阵 $A = [a_{ij}]$，记 $C = [C_{ij}]_{n\times n}$，则 C^{T} 为 A 的**伴随矩阵**，且 $C^{\mathrm{T}}A = \det(A)I_n$。

1.5 特征值和特征向量