线性代数入门

梁鑫　田根　杨一龙　编著

蔚　辉

清华大学丘成桐数学科学中心

报告中均非完整定义或准确表述。具体细节请参考教材。

Definition

集合 X 到 Y 的映射

$$\begin{array}{rrcl} f: & X & \to & Y, \\ & x & \mapsto & y = f(x). \end{array}$$

三种特殊情形：

1. 单射：零空间/核 $\mathcal{N}(f) = 0$.
2. 满射：列空间/像集 $\mathcal{R}(f)$.
3. 双射：f 既是单射又是满射。

Definition

集合 X 到 Y 的映射

$$\begin{aligned} f: \quad X &\to \quad Y, \\ x &\mapsto y = f(x). \end{aligned}$$

三种特殊情形:

1. 单射:零空间/核 $\mathcal{N}(f) = 0$.
2. 满射:列空间/像集 $\mathcal{R}(f)$.
3. 双射:f 既是单射又是满射。

- $X = Y$: 变换。
- 映射的复合: $g \circ f(x) = g(f(x))$。
- 逆映射: $f^{-1}f(x) = x \quad \forall x \in X$。

1. 数域 $\mathbb{F} \subset \mathbb{C}$.
2. 八条运算法则。
3. 子空间：对加法和数乘封闭。
4. $\mathcal{R}(A) \perp \mathcal{N}(A^{\mathrm{T}})$, $\mathcal{N}(A) = \mathcal{N}(A^{\mathrm{T}}A)$ 及其衍生等式。
5. $\mathcal{M}_1 \cap \mathcal{M}_2$, $\mathcal{M}_1 + \mathcal{M}_2$, $\mathcal{M}_1 \bigoplus \mathcal{M}_2$.

设 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 和 $\boldsymbol{t}_1, \cdots, \boldsymbol{t}_n$ 分别是 n 维线性空间 $\mathcal{U}$ 的一组基。应有

$$\left(\boldsymbol{t}_1, \cdots, \boldsymbol{t}_n\right) = \left(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\right)T.$$

其中 $T \in \mathbb{F}^{m \times n}$ 是过渡矩阵，必定可逆。

- 基变换，坐标系变换，变量代换。
- 转化为一组标准正交基。

给定数域$\mathbb{F}$上的线性空间$\mathcal{U}$ 和$\mathcal{V}$. 考虑从$\mathcal{U}$ 到$\mathcal{V}$ 的映射f.

Definition

称f 为从$\mathcal{U}$ 到$\mathcal{V}$ 的线性映射，如果线性运算满足可加性和齐次性:

1. 任取$\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{U}$, 有$f(\boldsymbol{x}+\boldsymbol{x}') = f(\boldsymbol{x}) + f(\boldsymbol{x}')$;
2. 任取$\boldsymbol{x} \in \mathcal{U}, k \in \mathbb{F}$, 有$f(k\boldsymbol{x}) = kf(\boldsymbol{x})$,

- 表示矩阵与线性映射一一对应。

设$\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$ 和$\boldsymbol{i}_1, \cdots, \boldsymbol{i}_m$ 分别是n 维线性空间$\mathcal{U}$ 和m 维线性空间$\mathcal{V}$的一组基。

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n = \left(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\right)\vec{x}.$$

$$f(\boldsymbol{x}) = \Big(f(\boldsymbol{e}_1), \cdots, f(\boldsymbol{e}_n)\Big)\vec{x} = \Big(\boldsymbol{i}_1, \cdots, \boldsymbol{i}_m\Big)F\vec{x},$$

其中表示矩阵为$F \in \mathbb{F}^{m\times n}$.

- 抽象线性映射f 的核、像集、特征对等问题可以转化为表示矩阵 F 的对应问题。

取 $\mathbb{R}^n$ 和 $\mathbb{R}^m$ 的标准基，有$\mathbb{R}^n$到 $\mathbb{R}^m$ 的线性映射f:

$$f(\boldsymbol{x}) = A\boldsymbol{x},$$

其中$A = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n] \in \mathbb{R}^{m\times n}$, 且$f(\boldsymbol{e}_i) = \boldsymbol{a}_i, \quad 1 \le i \le n$.

- 初等矩阵：对换 P_{ij}, 倍乘 $E_{ii;k}$, 倍加 $E_{ji;k}$.
- 初等变换不改变矩阵的秩。
- 倍加变换不改变矩阵的行列式函数值。
- 任何矩阵都可通过初等行变换化为阶梯形或唯一的行简化阶梯形。
- 使用阶梯形或行简化阶梯形求解线性方程组和进行基扩充。
- 分块矩阵的设计和使用：分块初等矩阵。

Definition

任意$\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$:

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}=\sum_{i=1}^{N}x_iy_i.$$

- 向量长度 $\|\boldsymbol{x}\|=\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}\right)^{\frac{1}{2}}=\left(\sum_{i=1}^{N}x_i^2\right)^{\frac{1}{2}}$.
- $\|\boldsymbol{x}\|=0\iff\boldsymbol{x}=\mathbf{0}_n$.
- 非零向量$\boldsymbol{x},\boldsymbol{y}$ 正交: $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}=0$.

1. $\mathbb{R}^m$ 的标准基：$\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$.
2. 一组基 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s \to$ Gram-Schmidt 正交化
   $\to$ 一组标准正交基 $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_s$.

$$[\boldsymbol{a}_1 \ \cdots \ \boldsymbol{a}_s] = [\boldsymbol{q}_1 \ \cdots \ \boldsymbol{q}_s][\boldsymbol{r}_1 \ \cdots \ \boldsymbol{r}_s] = [\boldsymbol{q}_1 \ \cdots \ \boldsymbol{q}_s]\begin{bmatrix} r_{11} & \cdots & r_{1s} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & r_{ss} \end{bmatrix}$$

   按列对应相等：

$$\begin{aligned} \boldsymbol{a}_1 &= r_{11}\boldsymbol{q}_1, \\ \boldsymbol{a}_2 &= r_{12}\boldsymbol{q}_1 + r_{22}\boldsymbol{q}_2, \\ &\cdots \end{aligned}$$

   利用 $\boldsymbol{q}_i$ 的正交性和单位化，依次求解 $\boldsymbol{r}_i$ 和 $\boldsymbol{q}_i$.
3. 列不满秩情形下的Gram-Schmidt 正交化。

1. $Q \in \mathbb{R}^{m\times m}$: $\quad Q^T Q = I_m$.
2. 保距：$\|Q\boldsymbol{x}\| = \|\boldsymbol{x}\| \quad \forall \boldsymbol{x} \in \mathbb{R}^m$.
3. 保内积：$(Q\boldsymbol{x})^{\mathrm{T}}(Q\boldsymbol{y}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}$.
4. 旋转变换 R_θ 和反射变换 $H_{\boldsymbol{v}} = I_m - 2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}$ with $\|\boldsymbol{v}\| = 1$。
5. 列正交矩阵 Q_1，扩充出 Q_2，得到正交矩阵 Q, i.e., 子空间的一组标准正交基扩充成全空间 $\mathbb{R}^m$ 的一组标准正交基。

- $P_{\mathcal{M}}$: 投影到子空间$\mathcal{M}$ 的正交投影线性变换，或正交投影矩阵。
- $P_{\mathcal{M}}^2 = P_{\mathcal{M}}^{\mathrm{T}} = P_{\mathcal{M}}$.
- $\forall \boldsymbol{a} \in \mathbb{R}^m$, 有唯一分解:

$$\boldsymbol{a} = \boldsymbol{a}_1 + \boldsymbol{a}_2 \quad \text{with } \boldsymbol{a}_1 = P_{\mathcal{M}}\boldsymbol{a} \in \mathcal{M}, \quad \boldsymbol{a}_2 \in \mathcal{M}^{\perp}.$$

- $\mathcal{M} = \operatorname{span}(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_r)$, 则有 $P_{\mathcal{M}} = Q_r Q_r^{\mathrm{T}}$,

$$\boldsymbol{a}_1 = P_{\mathcal{M}}\boldsymbol{a} = Q_r Q_r^{\mathrm{T}}\boldsymbol{a} = \sum_{j=1}^{r} (\boldsymbol{q}_j^{\mathrm{T}}\boldsymbol{a})\boldsymbol{q}_j.$$

- $\boldsymbol{a}$ 到子空间$\mathcal{M}$ 的最短距离:

$$\|\boldsymbol{a} - P_{\mathcal{M}}\boldsymbol{a}\| = \min_{\tilde{\boldsymbol{a}} \in \mathcal{M}} \|\boldsymbol{a} - \tilde{\boldsymbol{a}}\|.$$

Definition

$\delta(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n) = \delta(A) = \det(A) = |A|.$

1. 多线性:
$\delta(\cdots, k\boldsymbol{a}_i + k'\boldsymbol{a}_i', \cdots) = k\delta(\cdots, \boldsymbol{a}_i, \cdots) + k'\delta(\cdots, \boldsymbol{a}_i', \cdots).$
2. 列反对称性:$\delta(\cdots, \boldsymbol{a}_i, \cdots, \boldsymbol{a}_j, \cdots) = \delta(\cdots, \boldsymbol{a}_j, \cdots, \boldsymbol{a}_i, \cdots)$
3. 单位化条件:$\delta(I_n) = 1.$

- 倍加变换不改变行列式。
- 上(下)三角(分块)矩阵的行列式。
- A 可逆。$\Longleftrightarrow \det(A) \neq 0.$

- $\det(A) = a_{1j}C_{1j} + \cdots + a_{nj}C_{nj}$.
- $\boldsymbol{a}_i^T \boldsymbol{c}_j = \begin{cases} \det(A) & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$ with $\boldsymbol{c}_j = [C_{1j} \ \cdots \ C_{nj}]^{\mathrm{T}}$.
- $A^{-1} = \dfrac{C^T}{\det(A)}$，其中伴随矩阵 $C^{\mathrm{T}} = [\boldsymbol{c}_1 \ \cdots \ \boldsymbol{c}_n]^{\mathrm{T}}$.
- Cramer 法则. $A\boldsymbol{x} = \boldsymbol{b}$的解若存在，则有

$$x_i = \frac{\det(B_i)}{\det(A)} \text{ for } i = 1, \ldots, n.$$

$$A\boldsymbol{x} = \lambda\boldsymbol{x} \iff (\lambda I_n - A)\boldsymbol{x} = \boldsymbol{0} \text{ 有非零解。}$$

1. λ 是特征多项式

   $$p_A(\lambda) = \det(\lambda I_n - A) = \lambda^n - \mathrm{trace}(A)\lambda^{n-1} + \cdots + (-1)^n \det(A)$$

   的n 个复根。

2. Hamilton-Cayley 定理：$p_A(A) = O$.

3. $\boldsymbol{x}$ 是特征子空间$\mathcal{N}(\lambda I_n - A)$ 中的任意非零向量。

4. 不同特征值对应的特征向量线性无关。

5. 特征值的代数重数和几何重数。

6. 单特征值，半单特征值和亏损特征值。

Definition

$\exists \sigma \in \mathbb{R}$, 和非零向量 $\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{y} \in \mathbb{R}^m$, s.t.

$$A\boldsymbol{x} = \sigma \boldsymbol{y}, \quad A^{\mathrm{T}}\boldsymbol{y} = \sigma \boldsymbol{x}.$$

A 有 $r = \mathrm{rank}(A)$ 个正奇异值，从大到小排序，和 AA^{T} 或 $A^{\mathrm{T}}A$ 的 r 个正特征值一一对应：

$$\sigma_i = \sqrt{\lambda_i(AA^{\mathrm{T}})} = \sqrt{\lambda_i(A^{\mathrm{T}}A)} \quad \text{for } i = 1, \ldots, r.$$

1. n 阶实矩阵A正定：$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} > 0 \quad \forall$ 非零$\boldsymbol{x} \in \mathbb{R}^n$.

2. 实对称矩阵A 正定的等价结论：
   - $\lambda_i(A) > 0$ for $i = 1, \dots, n$.
   - $\exists$ 可逆T, s.t. $A = TT^{\mathrm{T}}$.
   - $A = LDL^T$, 其中对角矩阵D 的对角元素都> 0，L 为单位下三角矩阵.
   - 顺序主子式都为正。
   - 顺序主子阵都正定。

3. 实对称矩阵的Cholesky 分解：$A = LL^{\mathrm{T}}$, 其中L 为下三角矩阵。

4. A 半正定的等价结论。

$A \in \mathbb{R}^{m\times n}$ with $m \geq n$, 存在 QR 分解:

$$A = QR = [Q_1 \quad Q_2] \begin{bmatrix} R_1 \\ O \end{bmatrix} = Q_1 R_1.$$

其中 Q_1 是 $m \times n$ 列正交矩阵, R_1 是 n 阶对角元素非负的上三角矩阵。当 A 可逆时, R_1 的对角元素可全部取为正。

1. Gram-Schmidt 正交化: $A = Q_1 R_1$ with $\mathrm{rank(A)} \leq \mathrm{n}$.
2. 扩充出 Q_2.
3. 补足零元素, 由 R_1 得到 R.
4. $\mathrm{rank}(A) = r$: 得到列空间 $\mathcal{R}(A)$ 的一组标准正交基 Q_r, 用于计算正交投影矩阵 $P_A = Q_r Q_r^{\mathrm{T}}$。

1. $A \in \mathbb{C}^{n\times n}$ 可对角化：$\exists$ 可逆 X 和对角 Λ, s.t. A 有谱分解 $A = X\Lambda X^{-1}$.
2. Λ: n 个特征值。X: 对应的 n 个线性无关的特征向量。
3. A 可对角化。$\Longleftrightarrow$ A 的特征值都半单。
4. 实对称矩阵的特征值都是实且半单。互异特征值对应的特征向量正交。
5. 实对称矩阵总是有谱分解：$A = Q\Lambda Q^T$，其中 Q正交.

1. $A = U\Sigma V^T = U_r\Sigma_r V_r^{\mathrm{T}} = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^{\mathrm{T}} + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^{\mathrm{T}}$.
   - 实对称矩阵的谱分解：$A^{\mathrm{T}}A = V\Lambda V^{\mathrm{T}}$ 或 $AA^{\mathrm{T}} = U\Lambda U^{\mathrm{T}}$.
   - 得到 Σ 和 $V = [V_r\ V_2]$
   - $U_r = AV_r^T\Sigma_r^{-1}$.
   - 扩充出 U_2, 得到 $U = [U_r\ U_2]$.
2. 四个子空间 $\mathcal{R}(A), \mathcal{N}(A), \mathcal{R}(A^{\mathrm{T}}), \mathcal{N}(A^{\mathrm{T}})$ 与 U 和 V的关联。
3. Moore-Penrose 广义逆：$A^+ = V_r\Sigma_r^{-1}U_r^{\mathrm{T}}$.

- 取 $\mathcal{M} = \mathcal{R}(A)$, 记 P_A 为投影到列空间 $\mathcal{R}(A)$ 的正交投影矩阵: $\forall \boldsymbol{b} \in \mathbb{R}^m$, 有唯一分解:
$$\boldsymbol{b} = \boldsymbol{b}_1 + \boldsymbol{b}_2 = P_A \boldsymbol{b} + P_{A^\perp} \boldsymbol{b}.$$
其中, $\boldsymbol{b}_1 = P_A \boldsymbol{b} \in \mathcal{R}(A), \quad \boldsymbol{b}_1^{\mathrm{T}} \boldsymbol{b}_2 = 0.$

-
$$\|\boldsymbol{b} - P_A \boldsymbol{b}\| = \min_{\tilde{\boldsymbol{b}} \in \mathcal{R}(A)} \|\boldsymbol{b} - \tilde{\boldsymbol{b}}\|.$$

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{b} - A\boldsymbol{x}\| \iff A\boldsymbol{x} = P_A\boldsymbol{b}.$$ 总是有解，但不一定唯一。

- 正则化方法。

$$\boldsymbol{x} \text{ 是最小二乘解。} \iff A^{\mathrm{T}}A\boldsymbol{x} = A^{\mathrm{T}}\boldsymbol{b}. \text{（解可能不唯一）}$$

- QR 分解。

• $\operatorname{rank}(A) = n$，即 A 列满秩时,

$$\boldsymbol{x} \text{ 是最小二乘解。} \iff R_1\boldsymbol{x} = Q_1^{\mathrm{T}}\boldsymbol{b}.$$

• $\operatorname{rank}(A) = r < n$:

$$P_A = U_rU_r^{\mathrm{T}} \Rightarrow U_r\Sigma_rV_r^{\mathrm{T}}\boldsymbol{x} = U_rU_r^{\mathrm{T}}\boldsymbol{b}. \text{（解不唯一）}$$

最小二乘解中长度最小的解: $\boldsymbol{x} = V_r\Sigma_r^{-1}U_r^{\mathrm{T}}\boldsymbol{b} = A^{+}\boldsymbol{b}.$

存在对应的可逆矩阵。

1. 左相抵：$A = XB = PLU$.
2. 相抵：$A = XBY = XD_rY$.
3. 相似：$A = XBX^{-1} = XTX^{-1} = XJX^{-1} \in \mathbb{C}^{n\times n}$. J 为Jordan 块构成的对角分块矩阵。
4. 正交相似：$A = QBQ^{\mathrm{T}}$，其中 Q 正交. 实对称 $A = Q\Lambda Q^{\mathrm{T}}$.
5. 合同：$A = XTX^{\mathrm{T}} = XJX^{\mathrm{T}}$. Sylvester 惯性定律。
6. 对应的标准形和不变量。

给定非零$\boldsymbol{x}$, 其关于方阵A 的Rayleigh 商为$\dfrac{\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}$.

实对称A:

- 做特征值排序$\lambda_1 \geq \cdots \geq \lambda_n$,
- 取其正交单位特征向量组$\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$. 有:

$$\lambda_1 = \max_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}; \quad \lambda_i = \max_{\substack{\boldsymbol{x} \neq \boldsymbol{0} \\ \boldsymbol{x} \perp \operatorname{span}(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_{i-1})}} \frac{\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}} \quad \text{for } i = 2, \ldots, n.$$

$$\lambda_n = \min_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}; \quad \lambda_i = \min_{\substack{\boldsymbol{x} \neq \boldsymbol{0} \\ \boldsymbol{x} \in \operatorname{span}(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_{i-1})}} \frac{\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}} \quad \text{for } i = 1, \ldots, n-1.$$

1.

$$
\begin{aligned}
\|A\| &= \max_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = \max_{\|\boldsymbol{x}\|=1} \|A\boldsymbol{x}\| \\
&= \sqrt{\lambda_{\max}(A^{\mathrm{T}}A)} = \sqrt{\lambda_{\max}(AA^{\mathrm{T}})} \\
&= \sigma_{\max})(A) = \sigma_1.
\end{aligned}
$$

2. 低秩逼近。

$$
\|A - A_k\| = \min_{\mathrm{rank}(B)\leq k} \|A - B\| = \sigma_{k+1},
$$

其中 $k < r$, $A_k = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^{\mathrm{T}} + \cdots + \sigma_k \boldsymbol{u}_k \boldsymbol{v}_k^{\mathrm{T}}$.