§1 线性空间与子空间

程光辉

2020 年 2 月 28 日

定义 1 若 $\mathbf{P}$ 是一个数集, 其中任意两个数作加、减、乘、除 (除数不为 $\mathbf{0}$) 四则运算后, 其结果仍在 $\mathbf{P}$ 中, 就称 $\mathbf{P}$ 是一个数域.

常用的数域: 有理数域 $\mathbf{Q}$、实数域 $\mathbf{R}$、复数域 $\mathbf{C}$.

其它数域，例如：

$$Q(\sqrt{2})=\{a+b\sqrt{2}|a,b\in\mathbf{Q}\}.$$

定义 2 设 V 是一个非空集合，$\mathbf{P}$ 是一个数域. 在集合 V 的元素之间定义了一种代数运算，叫做加法：即对于 V 中的任意两个元素 α,β, 经过加法运算后在 V 中有唯一的元素 γ 与之对应，称 γ 为 α 和 β 的和，记 $\gamma=\alpha+\beta$. 在数域 $\mathbf{P}$ 与集合 V 的元素之间还定义了一种运算，叫做数量乘法：对于数域 $\mathbf{P}$ 的任一数 k 和集合 V 中任一元 α, 在 V 中有唯一的元素 δ 与它们对应，称 δ 为 k 与 α 的数量乘积，记 $\delta=k\alpha$. 如果加法和数量乘法还满足以下八条规则 (假设 $\alpha,\beta,v\in V$, $k,l\in\mathbf{P}$):

(1) $\alpha+\beta=\beta+\alpha$

(2) $(\alpha+\beta)+v=\alpha+(\beta+v)$

(3) 存在零元 $\mathbf{0}$，对 $\mathbf{V}$ 中任一元素 α, 有 $\alpha+\mathbf{0}=\alpha$ (加法零元)

(4) 对 V 中任一元素 α, 都存在 V 中元素 β, 使得 $\alpha+\beta=\mathbf{0}$(负元素)

(5) $1\alpha=\alpha$

(6) $k(l\alpha)=(kl)\alpha$

(7) $(k+l)\alpha=k\alpha+l\alpha$

(8) $k(\alpha+\beta)=k\alpha+k\beta$

则称 V 为数域 $\mathbf{P}$ 上的线性空间，其中元素也常称为向量.

例 1 判断下列描述的集合是否是线性空间:

(1) 所有 $m \times n$ 的实矩阵集合, 在实数域 $\mathbf{R}$ 上是否构成线性空间?

(2) 数域 $\mathbf{P}$ 按照本身的加法和数乘?

(3) 三维空间中所有不平行于某一条直线的向量, 按照向量的加法和数乘构成的集合?

(4) 全体实函数按照函数的加法与数乘构成的集合, 在实数域 $\mathbf{R}$ 上是否构成线性空间?

定义 3 在线性空间 V 中, 如果有 n 个向量 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 线性无关, 而 V 中任意 $n+1$ 个向量线性相关, 则称 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 为 V 的一组基底. n 称为线性空间 V 的维数, 常记为 $dimV = n$.

在线性空间 V 中, 由向量组 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 生成的子空间记为

$$L(\alpha_1, \alpha_2, \ldots, \alpha_s) \text{ 或 } \mathrm{span}\{\alpha_1, \alpha_2, \ldots, \alpha_s\},$$

其维数等于 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 的秩, $\alpha_1, \alpha_2, \ldots, \alpha_s$ 的最大无关组为子空间的基.

例 2 构造下列线性空间的一组基, 确定其维数:

(1) $\mathbf{P}^{m\times n}$?

(2) $\mathbf{P}^{n\times n}$ 上的所有对称矩阵?

(3) 复数域 $\mathbf{C}$ 在自身以及在实数域 $\mathbf{R}$ 上?

解: (1) $\mathbf{P}^{m\times n}$ 的基为 $E_{ij}, i = 1, 2, \cdots, m, j = 1, 2, \cdots, n$, E_{ij} 表示 (i, j) 位置为 1, 其余为 0 的 $m \times n$ 矩阵. 因此, $dim\mathbf{P}^{m\times n} = mn$.

(2) $\mathbf{P}^{n\times n}$ 上的所有对称矩阵的基为 $E_{ij} + E_{ji}, i, j = 1, 2, \cdots, n$, 其维数为 $\frac{n(n+1)}{2}$.

(3) 复数域 $\mathbf{C}$ 在自身的基为 1, 维数为 1; 复数域 $\mathbf{C}$ 在实数域 $\mathbf{R}$ 上的基为 $1, i$, 维数为 2.

定义 4 如果数域 $\mathbf{P}$ 上的线性空间 V 的一个非空子集 W 对于的两种运算也构成线性空间, 则称 W 为 V 的一个线性子空间 (简称子空间).

两个平凡子空间: V 自身与 $\{0\}$.

例 3 回答下列问题, 在实数域 $\mathbf{R}$ 上,

(1) $\mathbf{R}^3$ 的子空间有几类?

(2) $W = \left\{ \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} | a, b \in \mathbf{R} \right\}$ 是 $\mathbf{R}^{2\times 2}$ 的子空间吗?

(3) $\boldsymbol{A} \in \mathbf{R}^{\boldsymbol{n}\times\boldsymbol{n}}$, 则 $\boldsymbol{Ax}=\mathbf{0}$ 的所有解向量?

(4) $\mathbf{R}^2$ 是 $\mathbf{R}^3$ 的子空间吗?

解: (1) 4 类，分别是 $\mathbf{0}$ 维、$\mathbf{1}$ 维、$\mathbf{2}$ 维、$\mathbf{3}$ 维子空间.

(2) 是 $\mathbf{2}$ 维子空间.

(3) 是 $\boldsymbol{n}-\mathrm{rank}(\boldsymbol{A})$ 维子空间.

(4) 不是.

例 4 对矩阵 $\boldsymbol{A} \in \mathbf{R}^{\boldsymbol{m}\times\boldsymbol{n}}$, 有如下四个基本子空间:

(1) 值域空间 $\mathbf{R}(\boldsymbol{A}) = \{\boldsymbol{Ax}|\boldsymbol{x} \in \mathbf{R}^{\boldsymbol{n}}\} \subseteq \mathbf{R}^{\boldsymbol{m}}$ 列向量空间;

(2) 行向量空间 $\mathbf{R}(\boldsymbol{A}^{\boldsymbol{T}}) = \{\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{y}|\boldsymbol{y} \in \mathbf{R}^{\boldsymbol{m}}\} \subseteq \mathbf{R}^{\boldsymbol{n}}$;

(3) 核空间 $\mathbf{N}(\boldsymbol{A}) = \{\boldsymbol{x}|\boldsymbol{Ax}=\mathbf{0}\} \subseteq \mathbf{R}^{\boldsymbol{n}}$;

(4) 左核 $\mathbf{N}(\boldsymbol{A}^{\boldsymbol{T}}) = \{\boldsymbol{y}|\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{y}=\mathbf{0}\} \subseteq \mathbf{R}^{\boldsymbol{m}}$.

定理 1 对矩阵 $\boldsymbol{A} \in \mathbf{R}^{\boldsymbol{m}\times\boldsymbol{n}}$ 的四个基本子空间, 有如下性质:

(1) $\forall \boldsymbol{b} \in \mathbf{R}(\boldsymbol{A})$, 则 $\exists \boldsymbol{x} \in \mathbf{R}^{\boldsymbol{n}}$, *s.t.* $\boldsymbol{Ax}=\boldsymbol{b}$;

(2) $\forall \boldsymbol{a} \in \mathbf{R}(\boldsymbol{A}^{\boldsymbol{T}})$, 则 $\exists \boldsymbol{y} \in \mathbf{R}^{\boldsymbol{m}}$, *s.t.* $\boldsymbol{y}^{\boldsymbol{T}}\boldsymbol{A}=\boldsymbol{a}^{\boldsymbol{T}}$;

(3) $\mathbf{N}(\boldsymbol{A}) = \{\mathbf{0}\} \Leftrightarrow rank(\boldsymbol{A}) = \boldsymbol{n}$;

(4) $\mathbf{N}(\boldsymbol{A}^{\boldsymbol{T}}) = \{\mathbf{0}\} \Leftrightarrow rank(\boldsymbol{A}) = \boldsymbol{m}$.

例 5 设 $\boldsymbol{V_1}, \boldsymbol{V_2}$ 是线性空间 $\boldsymbol{V}$ 的两个非平凡子空间，证明：存在向量 $\boldsymbol{\alpha} \in \boldsymbol{V}$，使 $\boldsymbol{\alpha} \notin \boldsymbol{V_1}$, $\boldsymbol{\alpha} \notin \boldsymbol{V_2}$ 同时成立.

证明: 分 3 种情况进行讨论.

(1) 因为 $\boldsymbol{V_1}$ 是非平凡子空间，则存在向量 $\boldsymbol{\alpha} \notin \boldsymbol{V_1}$，如果 $\boldsymbol{\alpha} \notin \boldsymbol{V_2}$，则结论成立;

(2) 同理，$\boldsymbol{V_2}$ 是非平凡子空间，则存在向量 $\boldsymbol{\beta} \notin \boldsymbol{V_2}$，如果 $\boldsymbol{\beta} \notin \boldsymbol{V_1}$，则结论成立;

(3) 如果 $\boldsymbol{\alpha} \in \boldsymbol{V_2}$，就有 $\boldsymbol{\alpha} \in \boldsymbol{V_2}$，$\boldsymbol{\beta} \notin \boldsymbol{V_2}$；如果 $\boldsymbol{\beta} \in \boldsymbol{V_1}$，就有 $\boldsymbol{\alpha} \notin \boldsymbol{V_1}$，$\boldsymbol{\beta} \in \boldsymbol{V_1}$，这时有

$$\boldsymbol{\gamma} = \boldsymbol{\alpha} + \boldsymbol{\beta} \notin \boldsymbol{V_1}, \text{且} \boldsymbol{\gamma} = \boldsymbol{\alpha} + \boldsymbol{\beta} \notin \boldsymbol{V_2}.$$

综上，得证.

定义 5 设向量 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_p$ 是线性空间 $\boldsymbol{V}$ 的一组基, 对任一向量 $\boldsymbol{v} \in \boldsymbol{V}$, 如果存在一组标量 $c_1, c_2, \cdots, c_p$, 有 $\boldsymbol{v} = c_1\boldsymbol{\varepsilon}_1 + c_2\boldsymbol{\varepsilon}_2 + \cdots + c_p\boldsymbol{\varepsilon}_p$, 则称标量 $c_1, c_2, \cdots, c_p$ 为向量 $\boldsymbol{v}$ 在基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_p$ 下的坐标 *(coordinates)*, 称基向量 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_p$ 组成了向量表示的坐标系. 若记 $\boldsymbol{c} = (c_1, c_2, \cdots, c_p)^T$, 称向量 $\boldsymbol{c}$ 为向量 $\boldsymbol{v}$ 在基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_p$ 下的坐标向量.

定理 2 令 $\boldsymbol{V}$ 是一向量空间, $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_p$ 是空间 $\boldsymbol{V}$ 的一组基, 则对于任一向量 $\boldsymbol{v} \in \boldsymbol{V}$, 存在唯一的一组标量 $c_1, c_2, \cdots, c_p$, 使得 $\boldsymbol{v} = c_1\boldsymbol{\varepsilon}_1 + c_2\boldsymbol{\varepsilon}_2 + \cdots + c_p\boldsymbol{\varepsilon}_p$, 即向量在同一基下坐标唯一.

§2 空间分解与维数定理

程光辉

2020 年 3 月 3 日

定义 1 设 $\boldsymbol{V_1, V_2}$ 是线性空间 $\boldsymbol{V}$ 的子空间, 则称

$$\boldsymbol{V_1 + V_2 = \{\alpha_1 + \alpha_2 | \forall \alpha_1 \in V_1, \forall \alpha_2 \in V_2\}}$$

是子空间 $\boldsymbol{V_1}$ 与 $\boldsymbol{V_2}$ 的和, 而

$$\boldsymbol{V_1 \cap V_2 = \{\alpha | \alpha \in V_1 \text{且} \alpha \in V_2\}}$$

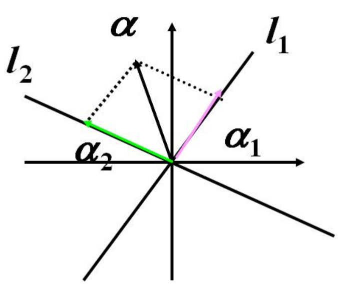
是子空间 $\boldsymbol{V_1}$ 与 $\boldsymbol{V_2}$ 的交.

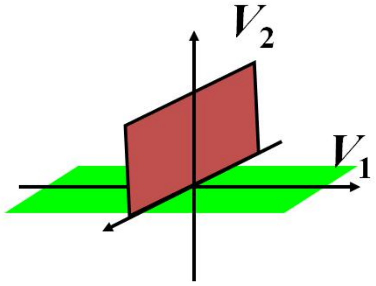
定理 1 设 $\boldsymbol{V_1, V_2}$ 是线性空间 $\boldsymbol{V}$ 的子空间, 则 $\boldsymbol{V_1 + V_2}$ 和 $\boldsymbol{V_1 \cap V_2}$ 也是 $\boldsymbol{V}$ 的子空间.

推论 1 设 $\boldsymbol{V_1, V_2}$ 是线性空间 $\boldsymbol{V}$ 的子空间, 则 $\boldsymbol{V_1 + V_2}$ 是 $\boldsymbol{V}$ 中包含子空间 $\boldsymbol{V_1}$ 和 $\boldsymbol{V_2}$ 的最小子空间.

推论 2 设 $\boldsymbol{V_1, V_2}$ 是线性空间 $\boldsymbol{V}$ 的子空间, 则 $\boldsymbol{V_1 \cap V_2}$ 是 $\boldsymbol{V}$ 中即包含于 $\boldsymbol{V_1}$, 又包含于 $\boldsymbol{V_2}$ 的最大子空间.

注意: $\boldsymbol{V_1 + V_2}$ 是子空间, $\boldsymbol{V_1 \cap V_2}$ 是子空间, 那么 $\boldsymbol{V_1 \cup V_2}$ 是子空间吗?





定理 2 *(维数定理)* 设 $\boldsymbol{V_1, V_2}$ 是线性空间 $\boldsymbol{V}$ 的子空间, 则:

$$dim(\boldsymbol{V_1}) + dim(\boldsymbol{V_2}) = dim(\boldsymbol{V_1 + V_2}) + dim(\boldsymbol{V_1 \cap V_2}).$$

证明：设 $\boldsymbol{V_1}, \boldsymbol{V_2}$ 的维数分别是 $\boldsymbol{s}, \boldsymbol{t}$，$\boldsymbol{V_1} \cap \boldsymbol{V_2}$ 的维数是 $\boldsymbol{m}$. 取 $\boldsymbol{V_1} \cap \boldsymbol{V_2}$ 的一组基

$$\boldsymbol{\alpha_1, \cdots, \alpha_m}.$$

如果 $\boldsymbol{m = 0}$, 这组基为空集，下面讨论中 $\boldsymbol{\alpha_1, \cdots, \alpha_m}$ 不出现，但不影响后面分析.

在 $\boldsymbol{\alpha_1, \cdots, \alpha_m}$ 基础上扩充成 $\boldsymbol{V_1}$ 的一组基

$$\boldsymbol{\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_{s-m}},$$

同理也扩充成 $\boldsymbol{V_2}$ 的一组基

$$\boldsymbol{\alpha_1, \cdots, \alpha_m, \gamma_1, \cdots, \gamma_{t-m}}.$$

下面证明向量组

$$\boldsymbol{\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_{s-m}, \gamma_1, \cdots, \gamma_{t-m}}$$

是 $\boldsymbol{V_1 + V_2}$ 的一组基. 即有维数定理成立.

因为

$$\begin{aligned} \boldsymbol{V_1} &= \operatorname{span}\{\boldsymbol{\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_{s-m}}\} \\ \boldsymbol{V_2} &= \operatorname{span}\{\boldsymbol{\alpha_1, \cdots, \alpha_m, \gamma_1, \cdots, \gamma_{t-m}}\}, \end{aligned}$$

所以

$$\boldsymbol{V_1 + V_2} = \operatorname{span}\{\boldsymbol{\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_{s-m}, \gamma_1, \cdots, \gamma_{t-m}}\}.$$

下面证明

$$\boldsymbol{\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_{s-m}, \gamma_1, \cdots, \gamma_{t-m}}$$

线性无关.

令

$$\boldsymbol{k_1\alpha_1 + \cdots + k_m\alpha_m + p_1\beta_1 + \cdots + p_{s-m}\beta_{s-m} + q_1\gamma_1 + \cdots + q_{t-m}\gamma_{t-m} = 0},$$

记

$$\boldsymbol{\alpha = k_1\alpha_1 + \cdots + k_m\alpha_m + p_1\beta_1 + \cdots + p_{s-m}\beta_{s-m} = -q_1\gamma_1 - \cdots - q_{t-m}\gamma_{t-m}}.$$

由 $\boldsymbol{\alpha = k_1\alpha_1 + \cdots + k_m\alpha_m + p_1\beta_1 + \cdots + p_{s-m}\beta_{s-m}}$ 知，$\boldsymbol{\alpha \in V_1}$；由 $\boldsymbol{\alpha = -q_1\gamma_1 - \cdots - q_{t-m}\gamma_{t-m}}$ 知，$\boldsymbol{\alpha \in V_2}$，于是有 $\boldsymbol{\alpha \in V_1 \cap V_2}$, 即 $\boldsymbol{\alpha}$ 可由 $\boldsymbol{\alpha_1, \cdots, \alpha_m}$ 线性表出. 令 $\boldsymbol{\alpha = l_1\alpha_1 + \cdots + l_m\alpha_m}$, 则

$$\boldsymbol{l_1\alpha_1 + \cdots + l_m\alpha_m + q_1\gamma_1 + \cdots + q_{t-m}\gamma_{t-m} = 0},$$

由于 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_{t-m}$ 线性无关，所以

$$l_1=\cdots=l_m=q_1=\cdots=q_{t-m}=0,$$

即 $\boldsymbol{\alpha}=\mathbf{0}$. 从而有

$$k_1\boldsymbol{\alpha}_1+\cdots+k_m\boldsymbol{\alpha}_m+p_1\boldsymbol{\beta}_1+\cdots+p_{s-m}\boldsymbol{\beta}_{s-m}=\mathbf{0}.$$

又因为 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{s-m}$ 线性无关，得

$$k_1=\cdots=k_m=p_1=\cdots=p_{s-m}=0.$$

综上，证明了

$$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{s-m},\boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_{t-m}$$

线性无关，因此它是 V_1+V_2 的一组基，故维数定理成立.

定义 2 设 V_1,V_2 是线性空间 V 的子空间, 若对 $\forall\boldsymbol{\alpha}\in V_1+V_2$, 有

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2\quad(\boldsymbol{\alpha}_1\in V_1,\boldsymbol{\alpha}_2\in V_2),$$

且这种表示是**唯一的**, 这个和 V_1+V_2 就称为直和, 记为 $V_1\oplus V_2$.

定理 3 设 V_1,V_2 是线性空间 V 的子空间, 则如下命题相互等价:

(1) V_1+V_2 是直和;

(2) $\mathbf{0}$ 向量表示法唯一, 即若 $\mathbf{0}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$ $(\boldsymbol{\alpha}_1\in V_1,\boldsymbol{\alpha}_2\in V_2)$, 必有 $\boldsymbol{\alpha}_1=\mathbf{0},\boldsymbol{\alpha}_2=\mathbf{0}$;

(3) $V_1\bigcap V_2=\{\mathbf{0}\}$.

推论 3 设 V_1,V_2 是线性空间 V 的子空间, 令 $W=V_1+V_2$, 则 $W=V_1\oplus V_2$ 充要条件是 $dim(W)=dim(V_1)+dim(V_2)$.

例 1 *(1)* 设 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 是线性空间 V 中线性无关的两个向量, 那么 $L(\boldsymbol{\alpha})+L(\boldsymbol{\beta})$ 是直和吗? $L(\boldsymbol{\alpha},\boldsymbol{\alpha}+\boldsymbol{\beta})+L(\boldsymbol{\beta})$ 呢? (其中 $L(\boldsymbol{\alpha},\boldsymbol{\beta})$ 代表由向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 生成的子空间.)

(2) 设 $\mathbf{R}^{n\times n}$ 表示所有 n 阶实方阵构成的线性空间, 而所有的实对称矩阵 $(A^T=A)$ 的集合 V_1 及所有实反对称矩阵 $(A^T=-A)$ 的集合 V_2, V_1 和 V_2 是子空间吗? 它们的和是直和吗? 如果是, $\mathbf{R}^{n\times n}=V_1\oplus V_2$ 吗?

解: (1) $L(\boldsymbol{\alpha})+L(\boldsymbol{\beta})$ 是直和; $L(\boldsymbol{\alpha},\boldsymbol{\alpha}+\boldsymbol{\beta})+L(\boldsymbol{\beta})$ 不是直和.

(2) 构成子空间; 是直和; $\mathbf{R}^{n\times n}=V_1\oplus V_2$ 成立.

多个子空间的直和是两个子空间直和的自然扩充, 分解式唯一.

定理 4 设 $\boldsymbol{V_1}, \boldsymbol{V_2}, \cdots, \boldsymbol{V_s}$ 是线性空间 $\boldsymbol{V}$ 的子空间, 则如下命题相互等价:

(1) $\boldsymbol{W} = \sum \boldsymbol{V}_i$ 是直和;

(2) $\mathbf{0}$ 向量表示法唯一;

(3) $\boldsymbol{V_i} \cap (\sum_{j \neq i} \boldsymbol{V_j}) = \{\mathbf{0}\}, \mathbf{1} \leq \boldsymbol{i} \leq \boldsymbol{s}$.

(4) $dim\boldsymbol{W} = \sum dim(\boldsymbol{V_i})$

§3 特征值和特征向量

程光辉

2020 年 3 月 14 日

1 线性变换

定义 1 若令 $\boldsymbol{V}$ 和 $\boldsymbol{W}$ 分别是 $\mathbf{C}^m$ 和 $\mathbf{C}^n$ 的子空间，则

$$\boldsymbol{T}:\ \boldsymbol{V} \to \boldsymbol{W}$$

称为子空间 $\boldsymbol{V}$ 到子空间 $\boldsymbol{W}$ 的变换 (或函数、映射)，它表示将子空间 $\boldsymbol{V}$ 的每一个向量变成子空间 $\boldsymbol{W}$ 的一个对应向量的一种规则. 即，对任一 $\boldsymbol{v} \in \boldsymbol{V}$，则存在 $\boldsymbol{w} \in \boldsymbol{W}$，有

$$\boldsymbol{w} = \boldsymbol{T}(\boldsymbol{v}),$$

并称子空间 $\boldsymbol{V}$ 是变换 $\boldsymbol{T}$ 的始集或定义域，称 $\boldsymbol{W}$ 是变换的终集或上域.

定义 2 若令 $\boldsymbol{V}$ 和 $\boldsymbol{W}$ 分别是 $\mathbf{C}^m$ 和 $\mathbf{C}^n$ 的子空间，并且 $\boldsymbol{T}:\ \boldsymbol{V} \to \boldsymbol{W}$ 是一种变换. 若对于 $\forall \boldsymbol{v_1}, \boldsymbol{v_2} \in \boldsymbol{V}$ 和任意标量 $\boldsymbol{c} \in \mathbf{C}$, 变换 $\boldsymbol{T}$ 满足叠加性

$$\boldsymbol{T}(\boldsymbol{v_1} + \boldsymbol{v_2}) = \boldsymbol{T}(\boldsymbol{v_1}) + \boldsymbol{T}(\boldsymbol{v_2})$$

和齐次性

$$\boldsymbol{T}(\boldsymbol{c}\boldsymbol{v_1}) = \boldsymbol{c}\boldsymbol{T}(\boldsymbol{v_1}),$$

则称 $\boldsymbol{T}$ 是线性变换或线性映射.

例 1 验证如下映射都是 $\mathbf{R}^3$ 上的线性变换:

(1) 坐标面上投影: $\boldsymbol{\sigma_{xy}} : \mathbf{R}^3 \to \mathbf{R}^3, (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \to (\boldsymbol{x}, \boldsymbol{y}, \mathbf{0})$;

(2) 关于坐标面的镜面反射: $\boldsymbol{\tau_{xy}} : \mathbf{R}^3 \to \mathbf{R}^3, (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \to (\boldsymbol{x}, \boldsymbol{y}, -\boldsymbol{z})$.

例 2 下面几个变换是否是线性变换?

(1) $\boldsymbol{T}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} y - z \\ x \\ y \end{bmatrix}$,

(2) $\boldsymbol{T}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+1 \\ z \\ y \end{bmatrix}$,

(3) $\boldsymbol{T}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} xy \\ -y \\ z \end{bmatrix}$,

(4) $\boldsymbol{T}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y \\ z \end{bmatrix}$.

若 T 是线性空间 $V_n(\mathrm{C})$ 的一种线性变换, $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是线性空间 $V_n(\mathrm{C})$ 的一组基底, 则

$$T(\varepsilon_j) = \sum_{i=1}^{n} a_{ij}\varepsilon_i = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)a_j, \quad j = 1, 2, \cdots, n,$$

其中 $a_j = (a_{1j}, a_{2j}, \cdots, a_{nj})^T \in \mathrm{C}^n$, 则

$$\Big(T(\varepsilon_1), T(\varepsilon_2), \cdots, T(\varepsilon_n)\Big) = \left(\sum_{i=1}^{n} a_{i1}\varepsilon_i, \sum_{i=1}^{n} a_{i2}\varepsilon_i, \cdots, \sum_{i=1}^{n} a_{in}\varepsilon_i\right) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)A,$$

其中 $A = (a_1, a_2, \cdots, a_n) \in \mathrm{C}^{n\times n}$, 称 A 是线性变换 T 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下对应的矩阵. 显然, 线性变换 T 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下和对应的矩阵 A 之间是一种 $1 \leftrightarrow 1$ 关系.

2 特征值和特征向量的概念

定义 3 设 T 是线性空间 $V_n(\mathrm{C})$ 的一种线性变换, 如果存在 $\lambda \in \mathrm{C}$ 和非零向量 $u \in V_n(\mathrm{C})$, 使 $T(u) = \lambda u$, 则 λ 叫做 T 的特征值, u 叫做 T 的属于特征值 λ 的特征向量.

若 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是线性空间 $V_n(\mathrm{C})$ 的一组基底, 对线性变换 T 的特征值 u, 有

$$u = \sum_{i=1}^{n} x_i\varepsilon_i = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)x,$$

其中 $x = (x_1, x_2, \cdots, x_n)^T \in \mathrm{C}^n$, 则

$$T(u) = \sum_{i=1}^{n} x_i T(\varepsilon_i) = \Big(T(\varepsilon_1), T(\varepsilon_2), \cdots, T(\varepsilon_n)\Big)x = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)Ax,$$

和

$$\lambda u = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)\lambda x,$$

于是有

$$Ax = \lambda x.$$

定义 4 设 $\boldsymbol{A} \in \mathbf{C}^{n\times n}$, 如果存在 $\boldsymbol{\lambda} \in \mathbf{C}$ 和非零向量 $\boldsymbol{u} \in \mathbf{C}^n$, 使 $\boldsymbol{Au} = \boldsymbol{\lambda u}$, 则 $\boldsymbol{\lambda}$ 叫做 $\boldsymbol{A}$ 的特征值, $\boldsymbol{u}$ 叫做 $\boldsymbol{A}$ 的属于特征值 $\boldsymbol{\lambda}$ 的特征向量.

从定义可以看出，使用矩阵 $\boldsymbol{A}$ 对向量 $\boldsymbol{u}$ 所作的线性变换 $\boldsymbol{Au}$ 不改变向量 $\boldsymbol{u}$ 的方向. 因此，线性变换是一种“保持方向不变”的变换. 为了确定向量 $\boldsymbol{u}$，可将定义式改写为

$$(\boldsymbol{\lambda E} - \boldsymbol{A})\boldsymbol{u} = \mathbf{0}.$$

由于要求 $\boldsymbol{u} \neq \mathbf{0}$，则矩阵 $\boldsymbol{\lambda E} - \boldsymbol{A}$ 奇异，即行列式

$$\det(\boldsymbol{\lambda E} - \boldsymbol{A}) = 0,$$

称矩阵 $(\boldsymbol{\lambda E} - \boldsymbol{A})$ 为 $\boldsymbol{A}$ 的特征矩阵，$\det(\boldsymbol{\lambda E} - \boldsymbol{A}) = 0$ 为 $\boldsymbol{A}$ 的特征方程，$\det(\boldsymbol{\lambda E} - \boldsymbol{A})$ 为 $\boldsymbol{A}$ 的特征多项式.

特征方程的根即为特征值.

定义 5 若 $\boldsymbol{\lambda}$ 是特征多项式 $\det(\boldsymbol{xE} - \boldsymbol{A}) = 0$ 的 $\boldsymbol{\mu}$ 重根，则称特征值 $\boldsymbol{\lambda}$ 的代数重复度为 $\boldsymbol{\mu}$; 若与特征值 $\boldsymbol{\lambda}$ 对应的线性无关特征向量个数为 γ，则称特征值 $\boldsymbol{\lambda}$ 的几何重复度为 γ.

特征值和特征向量的几个性质：

(1) $\det(\boldsymbol{A}) = \prod\limits_{i=1}^{n} \boldsymbol{\lambda}_i$;

(2) $\operatorname{tr}(\boldsymbol{A}) = \sum\limits_{i=1}^{n} \boldsymbol{\lambda}_i$;

(3) 同一特征值的几何重复度小于等于代数重复度;

(4) $\boldsymbol{A}$ 和 $\boldsymbol{A}^T$ 的特征值相同;

(5) 矩阵 $\boldsymbol{A} + \boldsymbol{\sigma E}$ 的特征值为 $\boldsymbol{\lambda} + \boldsymbol{\sigma}$;

(6) 不同特征值对应的特征向量线性无关.

3 矩阵相似

定义 6 设 $\boldsymbol{A}, \boldsymbol{B} \in \mathbf{C}^{n\times n}$, 如果存在可逆矩阵 $\boldsymbol{P} \in \mathbf{C}^{n\times n}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{BP} = \boldsymbol{A}$, 则称矩阵 $\boldsymbol{A}$ 相似于矩阵 $\boldsymbol{B}$.

定理 1 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征值.

定义 7 *Jordan* 标准型

$$\boldsymbol{J}=\begin{bmatrix}\boldsymbol{J}_{n_1} & \boldsymbol{O} & \cdots & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{J}_{n_2} & \cdots & \boldsymbol{O}\\ \vdots & \vdots & \ddots & \vdots\\ \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{J}_{n_r}\end{bmatrix}=\mathbf{diag}(\boldsymbol{J}_{n_1},\boldsymbol{J}_{n_2},\cdots,\boldsymbol{J}_{n_r}),$$

其中 $\boldsymbol{J}_{n_i}$ 为 $\boldsymbol{n_i}$ 阶对角元素都为 $\boldsymbol{\lambda_i}$, $\boldsymbol{\lambda_i}$ 正上方紧邻元素为 $\mathbf{1}$, 其余都为 $\mathbf{0}$ 的方阵.

定理 2 设 $\boldsymbol{A}\in\mathbf{C}^{\boldsymbol{n}\times\boldsymbol{n}}$ 有 $\boldsymbol{r}$ 个不同的特征值 $\boldsymbol{\lambda_1},\boldsymbol{\lambda_2},\cdots,\boldsymbol{\lambda_r}$, 其代数重数分别为 $\boldsymbol{n_1},\boldsymbol{n_2},\cdots,\boldsymbol{n_r}$, 则必存在可逆矩阵 $\boldsymbol{P}\in\mathbf{C}^{\boldsymbol{n}\times\boldsymbol{n}}$, 使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}=\mathbf{diag}\big(\boldsymbol{J}_{n_1}(\boldsymbol{\lambda_1}),\cdots,\boldsymbol{J}_{n_r}(\boldsymbol{\lambda_r})\big),$$

矩阵 $\boldsymbol{J}$ 叫做 $\boldsymbol{A}$ 的 *Jordan* 标准型.

§4 酉空间的分解与投影

程光辉

2020 年 3 月 13 日

1 欧式 (酉) 空间

定义 1 若 $\boldsymbol{V_n}(\mathbf{P})$ 上的映射 $(\boldsymbol{x},\boldsymbol{y})$: $\boldsymbol{V_n}(\mathbf{P})\times\boldsymbol{V_n}(\mathbf{P})\to\mathbf{P}$ 满足:

(1) $(\boldsymbol{x},\boldsymbol{x})\geq\mathbf{0}$; $(\boldsymbol{x},\boldsymbol{x})=\mathbf{0}$ 当且仅当 $\boldsymbol{x}=\mathbf{0}$; (正定性)

(2) $(\boldsymbol{x},\boldsymbol{y})=\overline{(\boldsymbol{y},\boldsymbol{x})}$, $\forall\boldsymbol{x},\boldsymbol{y}\in\boldsymbol{V_n}(\mathbf{P})$;((共轭) 对称性)

(3) $(\boldsymbol{x},\boldsymbol{\lambda y})=\boldsymbol{\lambda}(\boldsymbol{x},\boldsymbol{y})$, $\forall\boldsymbol{\lambda}\in\boldsymbol{P},\forall\boldsymbol{x},\boldsymbol{y}\in\boldsymbol{V_n}(\mathbf{P})$; (齐次性)

(4) $(\boldsymbol{x},\boldsymbol{y}+\boldsymbol{z})=(\boldsymbol{x},\boldsymbol{y})+(\boldsymbol{x},\boldsymbol{z})$, $\forall\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in\boldsymbol{V_n}(\mathbf{P})$, (可加性)

则称映射 $(\boldsymbol{x},\boldsymbol{y})$ 是 $\boldsymbol{V_n}(\mathbf{P})$ 上的内积; 若在 $\boldsymbol{n}$ 维线性空间 $\boldsymbol{V_n}(\mathbf{P})$ 中定义了内积, 则称该空间为内积空间.

例 1 *(1)* $\mathbf{C}^n$ 中的标准内积, $\forall\boldsymbol{x},\boldsymbol{y}\in\mathbf{C}^n$ 定义 $(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{x}^{\boldsymbol{H}}\boldsymbol{y}$;

(2) $\mathbf{C}^{m\times n}$ 中的标准内积, $\forall\boldsymbol{A},\boldsymbol{B}\in\mathbf{C}^{m\times n}$, 定义

$$(\boldsymbol{A},\boldsymbol{B})=tr(\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{B});$$

(3) 闭区间 $[\boldsymbol{a},\boldsymbol{b}]$ 全体连续函数构成的空间 $\boldsymbol{C}[\boldsymbol{a},\boldsymbol{b}]$ 中的内积, $\forall\boldsymbol{f},\boldsymbol{g}\in\boldsymbol{C}[\boldsymbol{a},\boldsymbol{b}]$, 定义

$$(\boldsymbol{f},\boldsymbol{g})=\int_a^b\overline{\boldsymbol{f}}\boldsymbol{g}d\boldsymbol{t};$$

(4) 若 $\boldsymbol{A}$ 为实对称正定矩阵, 则 $\mathbf{R}^n$ 可定义内积

$$(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{A}}=\sqrt{\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{y}},$$

称为 $\boldsymbol{A}$-内积.

定义 2 当 $\mathbf{P}=\mathbf{R}$ 时, $\boldsymbol{V_n}(\mathbf{R})$ 称为欧式空间; 当 $\mathbf{P}=\mathbf{C}$ 时, $\boldsymbol{V_n}(\mathbf{C})$ 称为酉空间.

定义 3 对任意 $\boldsymbol{x} \in \boldsymbol{V}_n(\mathbf{C})$, 称

$$\|\boldsymbol{x}\| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})}$$

为向量 $\boldsymbol{x}$ 的长度.

定理 1 向量长度的性质:

(1) $\|\boldsymbol{\lambda x}\| = |\boldsymbol{\lambda}|\|\boldsymbol{x}\|$;

(2) $\|\boldsymbol{x} - \boldsymbol{y}\|^2 + \|\boldsymbol{x} + \boldsymbol{y}\|^2 = 2\|\boldsymbol{x}\|^2 + 2\|\boldsymbol{y}\|^2$ *(平行四边形法则);*

(3) $|(\boldsymbol{x}, \boldsymbol{y})| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\|$ *(Cauchy 不等式);*

(4) $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ *(三角不等式).*

定义 4 两个非零向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 之间的夹角定义为

$$\cos\boldsymbol{\theta} = \frac{|(\boldsymbol{x}, \boldsymbol{y})|}{\sqrt{(\boldsymbol{x}, \boldsymbol{x})}\sqrt{(\boldsymbol{y}, \boldsymbol{y})}} = \frac{|\boldsymbol{x}^H\boldsymbol{y}|}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|}.$$

显然，当 $\boldsymbol{x}^H\boldsymbol{y} = 0$ 时, $\boldsymbol{\theta} = \frac{\pi}{2}$, 此时称向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 正交. 因此，两个常数向量正交定义如下.

定义 5 若两个常数向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的内积为零, 即 $(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^H\boldsymbol{y} = 0$, 则称它们是正交的, 并记为 $\boldsymbol{x} \perp \boldsymbol{y}$.

定义 6 若 $\boldsymbol{V}_1$ 和 $\boldsymbol{V}_2$ 是 $\boldsymbol{V}_n(\mathbf{C})$ 的两个子空间, 若 $\forall \boldsymbol{v}_1 \in \boldsymbol{V}_1$ 和 $\forall \boldsymbol{v}_2 \in \boldsymbol{V}_2$, 有 $(\boldsymbol{v}_1, \boldsymbol{v}_2) = 0$, 则称它们是正交的, 并记为 $\boldsymbol{V}_1 \perp \boldsymbol{V}_2$.

2 正交补子空间

定义 7 设 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，称 $\mathrm{N}(\boldsymbol{A}) = \{x|\boldsymbol{A}x = 0,\ x \in \mathbf{C}^n\}$ 为 $\boldsymbol{A}$ 的核 *(或零空间 **Null**)*, $\mathrm{R}(\boldsymbol{A}) = \{y|y = \boldsymbol{A}x,\ \forall x \in \mathbf{C}^n\}$ 为 $\boldsymbol{A}$ 的值域 *(**Range**)*.

定理 2 设 $\boldsymbol{A} \in \mathbf{C}^{n\times m}$，$\boldsymbol{B} \in \mathbf{C}^{n\times s}$，则 $\mathrm{R}(\boldsymbol{A}) \perp \mathrm{R}(\boldsymbol{B})$ 的充要条件是 $\boldsymbol{A}^H\boldsymbol{B} = \boldsymbol{O}$.

证明: 设 $\boldsymbol{A} = (\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m)$，$\boldsymbol{B} = (\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_s)$.

(必要性) 易知 $\boldsymbol{\alpha}_i \in \mathrm{R}(\boldsymbol{A})$，$i = 1, \cdots, m$ 和 $\boldsymbol{\beta}_j \in \mathrm{R}(\boldsymbol{B})$，$j = 1, \cdots, s$，因 $\mathrm{R}(\boldsymbol{A}) \perp R(\boldsymbol{B})$，有 $(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_j) = \boldsymbol{\alpha}_i^H\boldsymbol{\beta}_j = 0$，即 $\boldsymbol{A}^H\boldsymbol{B} = \boldsymbol{O}$.

(充分性) 对 $\forall \boldsymbol{y}_A \in \mathrm{R}(\boldsymbol{A})$ 和 $\forall \boldsymbol{y}_B \in \mathrm{R}(\boldsymbol{B})$，则存在向量 $\boldsymbol{x}_A \in \mathbf{C}^m$ 和 $\boldsymbol{x}_B \in \mathbf{C}^s$ 使得 $\boldsymbol{y}_A = \boldsymbol{A}\boldsymbol{x}_A$ 和 $\boldsymbol{y}_B = \boldsymbol{B}\boldsymbol{x}_B$.

$\boldsymbol{A}^H\boldsymbol{B} = \boldsymbol{O}$, 于是有

$$(\boldsymbol{y}_A, \boldsymbol{y}_B) = \boldsymbol{y}_A^H\boldsymbol{y}_B = \boldsymbol{x}_A^H\boldsymbol{A}^H\boldsymbol{B}\boldsymbol{x}_B = 0.$$

由于 $\boldsymbol{y}_A$ 和 $\boldsymbol{y}_B$ 的任意性，即 $\mathrm{R}(\boldsymbol{A}) \perp \mathrm{R}(\boldsymbol{B})$.

推论 1 设 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，则

$$\mathbf{R}(\boldsymbol{A}) \perp \mathbf{N}(\boldsymbol{A}^H); \quad \mathbf{N}(\boldsymbol{A}) \perp \mathbf{R}(\boldsymbol{A}^H).$$

推论 2 设 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，则

(1) $\dim \mathbf{R}(\boldsymbol{A}) + \dim \mathbf{N}(\boldsymbol{A}^H) = \boldsymbol{m}$;

(2) $\dim \mathbf{R}(\boldsymbol{A}^H) + \dim \mathbf{N}(\boldsymbol{A}) = \boldsymbol{n}$;

(3) $\mathbf{C}^m = \mathbf{R}(\boldsymbol{A}) \oplus \mathbf{N}(\boldsymbol{A}^H)$;

(4) $\mathbf{C}^n = \mathbf{R}(\boldsymbol{A}^H) \oplus \mathbf{N}(\boldsymbol{A})$.

定义 8 设酉空间 $\boldsymbol{V}_n(\mathrm{C})$ 的两个正交子空间 $\boldsymbol{V}_1$，$\boldsymbol{V}_2$，有 $\boldsymbol{V}_1 \perp \boldsymbol{V}_2$，且 $\boldsymbol{V}_1 + \boldsymbol{V}_2 = \boldsymbol{V}_n(\mathrm{C})$，则称 $\boldsymbol{V}_2$ 为 $\boldsymbol{V}_1$ 的正交补空间，记为 $\boldsymbol{V}_2 = \boldsymbol{V}_1^{\perp}$.

定理 3 酉空间 $\boldsymbol{V}_n(\mathrm{C})$ 的任意子空间 $\boldsymbol{V}_1$ 都有唯一的正交补.

证明： 若 $\boldsymbol{V}_1$ 是平凡子空间 $\{\mathbf{0}\}$ 或 $\boldsymbol{V}_n(\mathbf{C})$，则显然成立.

若 $\boldsymbol{V}_1$ 不是平凡子空间，取 $\boldsymbol{\epsilon}_1, \cdots, \boldsymbol{\epsilon}_m$ 为标准正交基，它可以扩充为 $\boldsymbol{V}_n(\mathbf{C})$ 的一组标准正交基

$$\boldsymbol{\epsilon}_1, \cdots, \boldsymbol{\epsilon}_m, \boldsymbol{\epsilon}_{m+1}, \cdots, \boldsymbol{\epsilon}_n.$$

即有 $\boldsymbol{V}_2 = \boldsymbol{L}(\boldsymbol{\epsilon}_{m+1}, \cdots, \boldsymbol{\epsilon}_n)$ 为 $\boldsymbol{V}_1$ 的正交补.

(唯一性) 设 $\boldsymbol{V}_2$，$\boldsymbol{V}_3$ 都是 $\boldsymbol{V}_1$ 的正交补，则

$$\boldsymbol{V}_n(\mathrm{C}) = \boldsymbol{V}_1 + \boldsymbol{V}_2, \quad \boldsymbol{V}_n(\mathrm{C}) = \boldsymbol{V}_1 + \boldsymbol{V}_3.$$

对 $\forall \boldsymbol{\alpha}_2 \in \boldsymbol{V}_2$，即有 $\boldsymbol{\alpha}_2 \in \boldsymbol{V}_n(\mathrm{C})$, 于是 $\boldsymbol{\alpha}_2 = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_3$, 其中 $\boldsymbol{\alpha}_1 \in \boldsymbol{V}_1$，$\boldsymbol{\alpha}_3 \in \boldsymbol{V}_3$. 又因为 $\boldsymbol{\alpha}_2 \perp \boldsymbol{\alpha}_1$ 和 $\boldsymbol{\alpha}_3 \perp \boldsymbol{\alpha}_1$，所以

$$\mathbf{0} = (\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1) = (\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_1) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1) + (\boldsymbol{\alpha}_3, \boldsymbol{\alpha}_1) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1),$$

即 $\boldsymbol{\alpha}_1 = \mathbf{0}$. 因此，$\boldsymbol{\alpha}_2 = \boldsymbol{\alpha}_3 \in \boldsymbol{V}_3$，即 $\boldsymbol{V}_2 \subseteq \boldsymbol{V}_3$.

同理可证 $\boldsymbol{V}_3 \subseteq \boldsymbol{V}_2$，即 $\boldsymbol{V}_2 = \boldsymbol{V}_3$. 唯一性得证.

3 投影与幂等矩阵

定义 9 设 $\boldsymbol{V}_n(\mathrm{C})$ 是线性空间，如果线性变换 $\boldsymbol{T}: \boldsymbol{V}_n(\mathrm{C}) \to \boldsymbol{V}_n(\mathrm{C})$ 具有 $\boldsymbol{T}^2 = \boldsymbol{T}$ 的性质，则称 $\boldsymbol{T}$ 是 $\boldsymbol{V}_n(\mathrm{C})$ 上的投影 *(也称投影算子或幂等算子)*.

定理 4 设 $\boldsymbol{T}$ 是 $\boldsymbol{V_n}(\mathrm{C})$ 上的投影，则

$$\boldsymbol{V_n}(\mathrm{C}) = \mathrm{R}(\boldsymbol{T}) \oplus \mathrm{N}(\boldsymbol{T}).$$

证明: 对 $\forall\boldsymbol{\alpha} \in \boldsymbol{V_n}(\mathrm{C})$，则有 $\boldsymbol{\alpha_1} = \boldsymbol{T}(\boldsymbol{\alpha})$，记 $\boldsymbol{\alpha_2} = \boldsymbol{\alpha} - \boldsymbol{\alpha_1}$. 因为 $\boldsymbol{T} = \boldsymbol{T^2}$，有

$$\boldsymbol{T}(\boldsymbol{\alpha_2}) = \boldsymbol{T}(\boldsymbol{\alpha} - \boldsymbol{\alpha_1}) = \boldsymbol{T}(\boldsymbol{\alpha}) - \boldsymbol{T}(\boldsymbol{\alpha_1}) = \boldsymbol{T}(\boldsymbol{\alpha}) - \boldsymbol{T^2}(\boldsymbol{\alpha}) = \mathbf{0}.$$

因此，$\boldsymbol{\alpha_2} \in \mathrm{N}(\boldsymbol{T})$. 所以 $\boldsymbol{V_n}(\mathrm{C}) = \mathrm{R}(\boldsymbol{T}) + \mathrm{N}(\boldsymbol{T})$.

$\forall\boldsymbol{\beta} \in \mathrm{R}(\boldsymbol{T}) \bigcap \mathrm{N}(\boldsymbol{T})$，则 $\boldsymbol{\beta} \in \mathrm{R}(\boldsymbol{T})$ 且 $\boldsymbol{\beta} \in \mathrm{N}(\boldsymbol{T})$，即有 $\exists\boldsymbol{\gamma} \in \boldsymbol{V_n}(\mathrm{C})$，使得 $\boldsymbol{\beta} = \boldsymbol{T}(\boldsymbol{\gamma})$ 和 $\boldsymbol{T}(\boldsymbol{\beta}) = \mathbf{0}$. 进一步可得

$$\boldsymbol{\beta} = \boldsymbol{T}(\boldsymbol{\gamma}) = \boldsymbol{T^2}(\boldsymbol{\gamma}) = \boldsymbol{T}(\boldsymbol{\beta}) = \mathbf{0}.$$

综上，有 $\boldsymbol{V_n}(\mathrm{C}) = \mathrm{R}(\boldsymbol{T}) \oplus \mathrm{N}(\boldsymbol{T})$.

定理 5 设 $\boldsymbol{V_n}(\mathrm{C}) = \boldsymbol{V_1} \oplus \boldsymbol{V_2}$，则存在投影 $\boldsymbol{T}$，使得

$$\mathrm{R}(\boldsymbol{T}) = \boldsymbol{V_1}, \quad \mathrm{N}(\boldsymbol{T}) = \boldsymbol{V_2}.$$

证明: 因为 $\boldsymbol{V_n}(\mathrm{C}) = \boldsymbol{V_1} \oplus \boldsymbol{V_2}$，对 $\forall\boldsymbol{\alpha} \in \boldsymbol{V_n}(\mathrm{C})$，则 $\boldsymbol{\alpha}$ 可以唯一的分解为 $\boldsymbol{\alpha} = \boldsymbol{\alpha_1} + \boldsymbol{\alpha_2}$，其中 $\boldsymbol{\alpha_1} \in \boldsymbol{V_1}$，$\boldsymbol{\alpha_2} \in \boldsymbol{V_2}$. 定义线性映射 $\boldsymbol{T}$ 满足 $\boldsymbol{T}(\boldsymbol{\alpha}) = \boldsymbol{\alpha_1}$，即

$$\boldsymbol{T}(\boldsymbol{\alpha}) = \boldsymbol{T}(\boldsymbol{\alpha_1} + \boldsymbol{\alpha_2}) = \boldsymbol{T}(\boldsymbol{\alpha_1}) + \boldsymbol{T}(\boldsymbol{\alpha_2}) = \boldsymbol{\alpha_1},$$

由于 $\boldsymbol{\alpha}$ 的任意性, 有 $\boldsymbol{T}(\boldsymbol{\alpha_1}) = \boldsymbol{\alpha_1}$, 进而可得 $\boldsymbol{T}(\boldsymbol{\alpha_2}) = \mathbf{0}$. 因此，$\boldsymbol{\alpha_2} \in \mathrm{N}(\boldsymbol{T})$. 由于 $\boldsymbol{\alpha}$ 的任意性和直和关系，有 $\mathrm{R}(\boldsymbol{T}) = \boldsymbol{V_1}$，$\mathrm{N}(\boldsymbol{T}) = \boldsymbol{V_2}$.

又因为

$$\boldsymbol{T^2}(\boldsymbol{\alpha}) = \boldsymbol{T}(\boldsymbol{\alpha_1}) = \boldsymbol{\alpha_1} = \boldsymbol{T}(\boldsymbol{\alpha}),$$

有 $\boldsymbol{T}$ 为投影.

4 正交投影

定义 10 设 $\boldsymbol{T}$ 是 $\boldsymbol{V_n}(\mathrm{C})$ 上的投影，$\boldsymbol{V_n}(\mathrm{C}) = \mathrm{R}(\boldsymbol{T}) \oplus \mathrm{N}(\boldsymbol{T})$. 如果 $\mathrm{R}^{\perp}(\boldsymbol{T}) = \mathrm{N}(\boldsymbol{T})$，则称 $\boldsymbol{T}$ 是正交投影.

定理 6 设 $\boldsymbol{A} \in \mathrm{C}^{n\times n}$，$\boldsymbol{A^2} = \boldsymbol{A}$，则 $\boldsymbol{A}$ 是正交投影的充分必要条件是 $\boldsymbol{A^H} = \boldsymbol{A}$.

证明: (充分性) 因为 $\boldsymbol{A^2} = \boldsymbol{A}$，知 $\mathrm{C}^n = \mathrm{R}(\boldsymbol{A}) \oplus \mathrm{N}(\boldsymbol{A})$. $\forall\boldsymbol{y} \in \mathrm{R}(\boldsymbol{A})$，$\forall\boldsymbol{x} \in \mathrm{N}(\boldsymbol{A})$，则 $\exists\boldsymbol{z} \in \mathrm{C}^n$ 使得 $\boldsymbol{y} = \boldsymbol{Az}$，$\boldsymbol{Ax} = \mathbf{0}$. 于是有

$$(\boldsymbol{y}, \boldsymbol{x}) = \boldsymbol{y^H x} = \boldsymbol{z^H A^H x} = \boldsymbol{z^H A x} = \mathbf{0},$$

即 $\boldsymbol{y} \perp \boldsymbol{x}$. 由于 $\boldsymbol{y},\ \boldsymbol{x}$ 的任意性，得 $\mathbf{R}^{\perp}(\boldsymbol{A}) = \mathbf{N}(\boldsymbol{A})$.

(必要性) 因为 $\boldsymbol{A}^2 = \boldsymbol{A}$，$\forall \boldsymbol{x} \in \mathbf{C}^n$，则 $\boldsymbol{A}\boldsymbol{x} \in \mathbf{R}(\boldsymbol{A})$，$\boldsymbol{x} - \boldsymbol{A}\boldsymbol{x} \in \mathbf{N}(\boldsymbol{A})$. 因为 $\mathbf{R}^{\perp}(\boldsymbol{A}) = \mathbf{N}(\boldsymbol{A})$，有

$$\mathbf{0} = (\boldsymbol{x} - \boldsymbol{A}\boldsymbol{x},\ \boldsymbol{A}\boldsymbol{x}) = \boldsymbol{x}^{\boldsymbol{H}}\boldsymbol{A}\boldsymbol{x} - \boldsymbol{x}^{\boldsymbol{H}}\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^{\boldsymbol{H}}(\boldsymbol{A} - \boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{A})\boldsymbol{x},$$

由 $\boldsymbol{x}$ 的任意性，知 $\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}}$ 为 Hermitian 矩阵，即 $\boldsymbol{A} = \boldsymbol{A}^{\boldsymbol{H}}$.

§5 初等矩阵

程光辉

2020 年 3 月 17 日

1 初等矩阵的一般形式

定义 1 设 u，$v \in \mathbf{C}^n$，$\sigma \in \mathbf{C}$，则形如

$$E(u, v; \sigma) = E_n - \sigma u v^H$$

的矩阵叫做初等矩阵，其中 E_n 为 n 阶单位矩阵.

若 $\sigma = 0$，有 $E(u, v; 0) = E_n$. 于是下面只讨论 $\sigma \neq 0$，u，v 非零向量情况的初等矩阵的性质.

定理 1 设 $u_1, u_2, \cdots, u_{n-1}$ 是 $v^\perp$ 的一组基，若 $u \in v^\perp$，则 $u_1, u_2, \cdots, u_{n-1}$ 是 $E(u, v; \sigma)$ 的 $n-1$ 个线性无关的特征向量；若 $u \notin v^\perp$，则 $u, u_1, u_2, \cdots, u_{n-1}$ 是 $E(u, v; \sigma)$ 的 n 个线性无关的特征向量.

证明： 因为

$$E(u, v; \sigma) u_i = E_n u_i - \sigma u v^H u_i = u_i,\ i = 1, 2, \cdots, n-1,$$

因此，u_i $(i = 1, 2, \cdots, n-1)$ 是 $E(u, v; \sigma)$ 的属于特征值 1 的线性无关特征向量.

若 $u \in v^\perp$，$E(u, v; \sigma)$ 存在 n 个线性无关的特征向量，则 v 一定是，即有

$$E(u, v; \sigma) v = E_n v - \sigma u v^H v = v - \sigma v^H v u = \lambda v,$$

则 u 和 v 共线 (必相关)，矛盾. 因此，则 $E(u, v; \sigma)$ 只有 $n-1$ 个线性无关的特征向量 $u_1, u_2, \cdots, u_{n-1}$.

若 $u \notin v^\perp$，有

$$E(u, v; \sigma) u = E_n u - \sigma u v^H u = (1 - \sigma v^H u) u,$$

则 u 是特征值 $1 - \sigma v^H u$ 对应的特征向量. 因此，则 $E(u, v; \sigma)$ 有 n 个线性无关的特征向量 $u, u_1, u_2, \cdots, u_{n-1}$.

推论 1 $E(u,v;\sigma)$ 的特征谱 *(即不考虑代数重数的特征值集合)* 为

$$\lambda\big(E(u,v;\sigma)\big)=\{1-\sigma v^H u,1,1,\cdots,1\}.$$

推论 2 当且仅当 $\sigma v^H u\neq 1$ 时，$E(u,v;\sigma)$ 可逆，且

$$E(u,v;\sigma)^{-1}=E(u,v;\frac{\sigma}{\sigma v^H u-1}).$$

推论 3 对任意的非零向量 $a,b\in \mathbf{C}^n$，存在 u,v,σ，使得

$$E(u,v;\sigma)a=b.$$

2 初等酉阵

定义 2 设 $u\in\mathbf{C}^n$，且 $u^H u=1$，则

$$H(u)=E(u,u;2)=E_n-2uu^H$$

称为初等酉阵，或 *Householder* 变换.

定理 2 $H(u)^H=H(u)=H(u)^{-1}$.

定理 3 $H(u)$ 是镜像变换. 即对 $\forall a\in u^{\perp}$，有

$$H(u)(a+ru)=a-ru,\quad r\in\mathbf{C},$$

也就是说 $H(u)$ 是关于 u 的垂直超平面的镜像.

3 酉变换与酉矩阵

定义 3 若线性空间 $V_n(\mathbf{C})$ 的变换 T 满足:

$$(T(x),T(y))=(x,y),\forall x,y\in V_n(\mathbf{C}),$$

则称 T 为 $V_n(\mathbf{C})$ 的酉变换.

酉变换及其对应的矩阵有非常好的性质, 在实际工程计算中有非常广泛的应用.

定理 4 设 T 是 $V_n(\mathbf{C})$ 的线性变换, 则下列命题等价:

(1) T 是酉变换;

(2) $\|T(x)\|=\|x\|,\forall x\in V_n(\mathbf{C})$;

(3) 设 $\varepsilon_1, \cdots, \varepsilon_n$ 是 $V_n(\mathbf{C})$ 的标准正交基, 则 $T(\varepsilon_1), \cdots, T(\varepsilon_n)$ 也是它的标准正交基;

(4) T 在任一标准正交基下的矩阵 A 是酉矩阵, 即 $A^H A = AA^H = E_n$.

定理 5 设 $A \in \mathbf{C}^{n\times n}$ 是酉矩阵, 则

(1) $(Ax, Ay) = (x, y)$;

(2) $\|Ax\| = \|x\|, \forall x \in \mathbf{C}^n$;

(3) A^H 也是酉矩阵;

(4) 若 $B \in \mathbf{C}^{n\times n}$ 是酉矩阵, 则 AB, BA 也是酉矩阵;

(5) 酉矩阵的特征值的模为 1.

§6 Kronecker 乘积

程光辉

2020 年 3 月 20 日

1 基本概念和基本性质

定义 1 设 $\boldsymbol{A}=(a_{ij})\in \mathrm{P}^{m\times n}, \boldsymbol{B}=(b_{ij})\in \mathrm{P}^{p\times q}$, 则

$$\boldsymbol{A}\otimes\boldsymbol{B}=\begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

称为矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的 *Kronecker* 乘积 *(*或直积*,* 张量积*)*.

例 1 若 $\boldsymbol{A}=\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}, \boldsymbol{B}=\begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}$, 则

$$\boldsymbol{A}\otimes\boldsymbol{B}=\begin{bmatrix} B & 2B & 3B \\ 3B & 2B & B \end{bmatrix}=\begin{bmatrix} 2 & 1 & 4 & 2 & 6 & 3 \\ 2 & 3 & 4 & 6 & 6 & 9 \\ 6 & 3 & 4 & 2 & 2 & 1 \\ 6 & 9 & 4 & 6 & 2 & 3 \end{bmatrix}.$$

一般情况下，$\boldsymbol{A}\otimes\boldsymbol{B}\neq\boldsymbol{B}\otimes\boldsymbol{A}$.

定理 1 设 $\boldsymbol{A}\in \mathrm{P}^{m\times n}, \boldsymbol{B}\in \mathrm{P}^{p\times q}, \boldsymbol{C}\in \mathrm{P}^{r\times s}, \boldsymbol{D}\in \mathrm{P}^{k\times h}$, 则

(1) 单位矩阵之积*:* $\boldsymbol{E}_m\otimes\boldsymbol{E}_n=\boldsymbol{E}_{mn}$*;*

(2) 纯量积*:* $\lambda(\boldsymbol{A}\otimes\boldsymbol{B})=(\lambda\boldsymbol{A})\otimes\boldsymbol{B}=\boldsymbol{A}\otimes(\lambda\boldsymbol{B}),\forall\lambda\in\mathrm{P}$*;*

(3) 分配律*:* 当 $m=p, n=q$ 时*,* $(\boldsymbol{A}+\boldsymbol{B})\otimes\boldsymbol{C}=(\boldsymbol{A}\otimes\boldsymbol{C})+(\boldsymbol{B}\otimes\boldsymbol{C})$*,* $\boldsymbol{C}\otimes(\boldsymbol{A}+\boldsymbol{B})=(\boldsymbol{C}\otimes\boldsymbol{A})+(\boldsymbol{C}\otimes\boldsymbol{B})$*;*

(4) 结合律*:* $(\boldsymbol{A}\otimes\boldsymbol{B})\otimes\boldsymbol{C}=\boldsymbol{A}\otimes(\boldsymbol{B}\otimes\boldsymbol{C})$*;*

(5) 转置及共轭: $(\boldsymbol{A}\otimes\boldsymbol{B})^{\boldsymbol{T}}=\boldsymbol{A}^{\boldsymbol{T}}\otimes\boldsymbol{B}^{\boldsymbol{T}}$, $\overline{(\boldsymbol{A}\otimes\boldsymbol{B})}=\bar{\boldsymbol{A}}\otimes\bar{\boldsymbol{B}}$;

(6) 混合积: 当 $\boldsymbol{n}=\boldsymbol{r},\boldsymbol{q}=\boldsymbol{k}$ 时, 有 $(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D})=(\boldsymbol{AC})\otimes(\boldsymbol{BD})$;

(7) 逆: 若 $\boldsymbol{A}^{-1},\boldsymbol{B}^{-1}$ 存在, 则

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

(8) 迹: 当 $\boldsymbol{m}=\boldsymbol{n},\boldsymbol{p}=\boldsymbol{q}$ 时, 有 $tr(\boldsymbol{A}\otimes\boldsymbol{B})=tr\boldsymbol{A}\cdot tr\boldsymbol{B}$;

(9) 秩: $rank(\boldsymbol{A}\otimes\boldsymbol{B})=rank\boldsymbol{A}\cdot rank\boldsymbol{B}$;

(10) 行列式: 当 $\boldsymbol{m}=\boldsymbol{n},\boldsymbol{p}=\boldsymbol{q}$ 时, 有 $\det(\boldsymbol{A}\otimes\boldsymbol{B})=(\det\boldsymbol{A})^{\boldsymbol{p}}(\det\boldsymbol{B})^{\boldsymbol{m}}$;

(11) 当 $\boldsymbol{A},\boldsymbol{B}$ 为对称矩阵时, $\boldsymbol{A}\otimes\boldsymbol{B}$ 也是对称矩阵; 当 $\boldsymbol{A},\boldsymbol{B}$ 为 *Herimtian* 矩阵时, $\boldsymbol{A}\otimes\boldsymbol{B}$ 也是 *Hermitian* 矩阵;

(12) $\boldsymbol{U}\in\mathbf{P}^{\boldsymbol{n}\times\boldsymbol{n}},\boldsymbol{V}\in\boldsymbol{P}^{\boldsymbol{m}\times\boldsymbol{m}}$ 均为酉矩阵, $\boldsymbol{U}\otimes\boldsymbol{V}$ 也是酉矩阵;

(13) 若令 $\boldsymbol{A}^{[0]}=\boldsymbol{1}$, $\boldsymbol{A}^{[1]}=\boldsymbol{A},\boldsymbol{A}^{[\boldsymbol{k}]}=\boldsymbol{A}\otimes\boldsymbol{A}\otimes\cdots\otimes\boldsymbol{A}$, 则 $(\boldsymbol{AB})^{[\boldsymbol{k}]}=\boldsymbol{A}^{[\boldsymbol{k}]}\boldsymbol{B}^{[\boldsymbol{k}]}$.

证明：(6)

$$\begin{aligned}(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D})&=\begin{bmatrix}a_{11}B&\cdots&a_{1n}B\\\vdots&\ddots&\vdots\\a_{m1}B&\cdots&a_{mn}B\end{bmatrix}\begin{bmatrix}c_{11}D&\cdots&c_{1s}D\\\vdots&\ddots&\vdots\\c_{n1}D&\cdots&c_{ns}D\end{bmatrix}\\&=\begin{bmatrix}\sum\limits_{k=1}^{n}a_{1k}c_{k1}BD&\cdots&\sum\limits_{k=1}^{n}a_{1k}c_{ks}BD\\\vdots&\ddots&\vdots\\\sum\limits_{k=1}^{n}a_{mk}c_{k1}BD&\cdots&\sum\limits_{k=1}^{n}a_{mk}c_{ks}BD\end{bmatrix}\\&=AC\otimes BD.\end{aligned}$$

(10) 由 Jordan 标准型分解知，存在可逆矩阵 $\boldsymbol{P},\boldsymbol{Q}$，使得

$$A=P^{-1}\begin{bmatrix}\lambda_1&&&\star\\&\lambda_2&&\\&&\ddots&\\0&&&\lambda_m\end{bmatrix}P=P^{-1}J_AP$$

和

$$B=Q^{-1}\begin{bmatrix}\mu_1&&&\star\\&\mu_2&&\\&&\ddots&\\0&&&\mu_m\end{bmatrix}Q=Q^{-1}J_BQ,$$

则利用第 (6) 条性质有

$$A \otimes B = (P^{-1}J_A P) \otimes (Q^{-1}J_B Q) = (P \otimes Q)^{-1}(J_A \otimes J_B)(P \otimes Q).$$

于是有

$$\begin{aligned}
\det(A \otimes B) &= \det(J_A \otimes J_B) \\
&= (\prod_{j=1}^{p} \lambda_1 \mu_j)(\prod_{j=1}^{p} \lambda_2 \mu_j) \cdots (\prod_{j=1}^{p} \lambda_m \mu_j) \\
&= (\prod_{i=1}^{m} \lambda_i)^p (\prod_{j=1}^{p} \mu_j)^m \\
&= (\det(A))^p (\det(B))^m.
\end{aligned}$$

2 Kronecker 积的特征值

定理 2 设 $\lambda_i(i = 1, 2, \ldots, m)$ 是 $A \in \mathrm{C}^{m\times m}$ 的特征值, x_i 是相应的特征向量; $\mu_j(j = 1, 2, \ldots, n)$ 是 $B \in \mathrm{C}^{n\times n}$ 的特征值, y_j 是相应的特征向量, 则 $A \otimes B$ 的 mn 个特征值是 $\lambda_i \mu_j$, 对应的特征向量是 $x_i \otimes y_j$, $1 \le i \le m, 1 \le j \le n$.

证明: 因为 $Ax_i = \lambda_i x_i$, $By_j = \mu_j y_j$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, 则

$$\begin{aligned}
(A \otimes B)(x_i \otimes y_j) &= Ax_i \otimes By_j \\
&= \lambda_i x_i \otimes \mu_j y_j \\
&= \lambda_i \mu_j (x_i \otimes y_j),
\end{aligned}$$

得证.

定义 2 m 阶矩阵 A 与 n 阶矩阵 B 的 *Kronecker* 和定义为

$$A \oplus_k B = A \otimes E_n + E_m \otimes B.$$

定理 3 设 $\lambda_i(i = 1, 2, \ldots, m)$ 是 $A \in \mathrm{C}^{m\times m}$ 的特征值, x_i 是相应的特征向量; $\mu_j(j = 1, 2, \ldots, n)$ 是 $B \in \mathrm{C}^{n\times n}$ 的特征值, y_j 是相应的特征向量, 则 $A \oplus_k B$ 的 mn 个特征值是 $\lambda_i + \mu_j$, 对应的特征向量是 $x_i \otimes y_j$, $1 \le i \le m, 1 \le j \le n$.

证明: 因为 $Ax_i = \lambda_i x_i$, $By_j = \mu_j y_j$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, 则

$$\begin{aligned}
(A \oplus_k B)(x_i \otimes y_j) &= (A \otimes E_n)(x_i \otimes y_j) + (E_m \otimes B)(x_i \otimes y_j) \\
&= (Ax_i) \otimes y_j + x_i \otimes (By_j) \\
&= (\lambda_i + \mu_j)(x_i \otimes y_j),
\end{aligned}$$

得证.

3 向量化算符

设矩阵

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

记 A 的列为 $A_1, A_2, \ldots, A_n$, 即 $A=(A_1, A_2, \ldots, A_n)$. 向量化算符: $\mathrm{Vec}(A)=\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix}$

定理 4 设 $A \in \mathbf{C}^{m\times n}$, $X \in \mathbf{C}^{n\times r}$, $B \in \mathbf{C}^{r\times s}$, 则

$$\mathrm{Vec}(AXB)=(B^T \otimes A)\mathrm{Vec}(X).$$

证明: 令 B_k 为矩阵 B 的第 k 列, 则

$$\begin{aligned} (AXB)_k &= A(XB)_k \\ &= AXB_k \\ &= A(X_1 b_{1k} + \ldots + X_r b_{rk}) \\ &= b_{1k}AX_1 + \ldots + b_{rk}AX_r \\ &= \begin{bmatrix} b_{1k}A, \ldots, b_{rk}A \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ X_r \end{bmatrix} \\ &= (B_k^T \otimes A)\mathrm{Vec}(X), \quad k=1,\ldots,s. \end{aligned}$$

于是有

$$\mathrm{Vec}(AXB)=\begin{bmatrix} B_1^T \otimes A \\ \vdots \\ B_s^T \otimes A \end{bmatrix} \mathrm{Vec}(X)=(B^T \otimes A)\mathrm{Vec}(X).$$

推论 1 设 $A \in \mathbf{C}^{m\times m}$, $X \in \mathbf{C}^{m\times n}$, $B \in \mathbf{C}^{n\times n}$, 则

(1) $\mathrm{Vec}(AX)=(E_n \otimes A)\mathrm{Vec}(X)$;

(2) $\mathrm{Vec}(XB)=(B^T \otimes E_m)\mathrm{Vec}(X)$;

(3) $\mathrm{Vec}(AX+XB)=(E_n \otimes A + B^T \otimes E_m)\mathrm{Vec}(X)$.

§1 向量范数

程光辉

2020 年 3 月 22 日

定义 1 设映射 $\|\cdot\|: \mathbf{C}^n \to \mathbf{R}$ 满足:

(1) 正定性 $\|\boldsymbol{x}\| \geq \mathbf{0}$，当且仅当 $\boldsymbol{x} = \mathbf{0}$ 时，$\|\boldsymbol{x}\| = \mathbf{0}$；

(2) 齐次性 $\|\boldsymbol{\lambda x}\| = |\boldsymbol{\lambda}|\|\boldsymbol{x}\|, \boldsymbol{\lambda} \in \mathbf{C}, \boldsymbol{x} \in \mathbf{C}^n$；

(3) 三角不等式 $\|\boldsymbol{x}+\boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$, $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbf{C}^n$,

则称映射 $\|\cdot\|$ 为 $\mathbf{C}^n$ 上向量 $\boldsymbol{x}$ 的范数.

向量范数的性质

(1) $\|\mathbf{0}\| = \mathbf{0}$

(2) $\boldsymbol{x} \neq \mathbf{0}$ 时，$\left\|\frac{\mathbf{1}}{\|\boldsymbol{x}\|}\boldsymbol{x}\right\| = \mathbf{1}$；

(3) 对任意 $\boldsymbol{x} \in \mathbf{C}^n$，有 $\|-\boldsymbol{x}\| = \|\boldsymbol{x}\|$；

(4) 对任意 $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{C}^n$，有 $|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|| \leq \|\boldsymbol{x}-\boldsymbol{y}\|$.

证明： (4) 因为

$$\|\boldsymbol{x}\| = \|(\boldsymbol{x}-\boldsymbol{y}) + \boldsymbol{y}\| \leq \|\boldsymbol{x}-\boldsymbol{y}\| + \|\boldsymbol{y}\|,$$

即有 $\|\boldsymbol{x}\| - \|\boldsymbol{y}\| \leq \|\boldsymbol{x}-\boldsymbol{y}\|$.

又因

$$\|\boldsymbol{y}\| - \|\boldsymbol{x}\| \leq \|\boldsymbol{y}-\boldsymbol{x}\| = \|\boldsymbol{x}-\boldsymbol{y}\|,$$

有 $-\|\boldsymbol{x}-\boldsymbol{y}\| \leq \|\boldsymbol{x}\| - \|\boldsymbol{y}\|$.

综上，$|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|| \leq \|\boldsymbol{x}-\boldsymbol{y}\|$ 得证.

例 1 设 $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n)^{\mathrm{T}} \in \mathbf{C}^n$，则

(1) $\|\boldsymbol{x}\|_1 = \sum\limits_{i=1}^{n} |\boldsymbol{x}_i| \Leftarrow$ *1*-范数

(2) $\|x\|_2 = \left(\sum\limits_{i=1}^{n} |x_i|^2\right)^{1/2} \Leftarrow$ *2-范数*

(3) $\|x\|_\infty = \max\limits_{1\le i\le n}\{|x_i|\} \Leftarrow$ *无穷范数*

证明：(2) 正定性和齐次性显然成立，下面只证明三角不等式.

设 $x=(x_1,x_2,\cdots,x_n)^T$，$y=(y_1,y_2,\cdots,y_n)^T$，则

$$\begin{aligned}|(x,y)|^2 &= |\overline{x}_1y_1+\overline{x}_2y_2+\cdots+\overline{x}_ny_n|^2\\ &\le \left(|x_1|^2+|x_2|^2+\cdots+|x_n|^2\right)\left(|y_1|^2+|y_2|^2+\cdots+|y_n|^2\right)\\ &= \|x\|_2^2\|y\|_2^2.\end{aligned}$$

于是

$$\begin{aligned}\|x+y\|_2^2 &= (x+y)^H(x+y)\\ &= x^Hx+x^Hy+y^Hx+y^Hy\\ &\le |x^Hx|+|x^Hy|+|y^Hx|+|y^Hy|\\ &\le \|x\|_2^2+2\|x\|_2\|y\|_2+\|y\|_2^2\\ &= (\|x\|_2+\|y\|_2)^2,\end{aligned}$$

即

$$\|x+y\|_2 \le \|x\|_2+\|y\|_2.$$

引理 1 *若 u 和 v 是非负实数，p 和 q 是正实数，且满足条件 $p,q>1$ 和 $\frac{1}{p}+\frac{1}{q}=1$，则恒有不等式*

$$uv \le \frac{1}{p}u^p+\frac{1}{q}v^q.$$

证明：根据定积分的几何意义知，以 u 和 v 为边长的矩形面积小于等于两个曲边梯形面积 $\int_0^u x^{p-1}dx$ 和 $\int_0^v y^{\frac{1}{p-1}}dy$ 之和，即

$$\begin{aligned}uv &\le \int_0^u x^{p-1}dx+\int_0^v y^{\frac{1}{p-1}}dy\\ &= \frac{1}{p}u^p+\int_0^v y^{\frac{q}{p}}dy\\ &= \frac{1}{p}u^p+(\frac{q}{p}+1)^{-1}v^{\frac{q}{p}+1}\\ &= \frac{1}{p}u^p+\frac{1}{q}v^q.\end{aligned}$$

定义 2 $\|x\|_p = \left(\sum\limits_{i=1}^{n}|x_i|^p\right)^{1/p}$，$1\le p<\infty$. *显然是一种映射关系.*

定理 1 *(Hölder 不等式)* 若 $p, q > 0$, 且 $\frac{1}{p}+\frac{1}{q}=1$, 则对 $\mathbf{C}^n$ 任意向量 $x=(x_1, x_2, \cdots, x_n)^T, y=(y_1, y_2, \cdots, y_n)^T$ 都有

$$\sum_{i=1}^{n}|x_i|\cdot|y_i| \leq \left(\sum_{i=1}^{n}|x_i|^p\right)^{1/p}\left(\sum_{i=1}^{n}|y_i|^q\right)^{1/q}$$

证明：若 x, y 至少有一个为零向量，则显然成立.

下面证明 x, y 均为非零情况. 令

$$u=\frac{|x_i|}{\|x\|_p}, \quad v=\frac{|y_i|}{\|y\|_q}.$$

由引理 1，有

$$\frac{|x_i||y_i|}{\|x\|_p\|y\|_q} \leq \frac{1}{p}\frac{|x_i|^p}{\|x\|_p^p}+\frac{1}{q}\frac{|y_i|^q}{\|y\|_q^q}, \quad (1 \leq i \leq n),$$

因此，对 i 求和得

$$\begin{aligned}\sum_{i=1}^{n}\frac{|x_i||y_i|}{\|x\|_p\|y\|_q} &\leq \frac{1}{p\|x\|_p^p}\sum_{i=1}^{n}|x_i|^p+\frac{1}{q\|y\|_q^q}\sum_{i=1}^{n}|y_i|^q \\ &= \frac{1}{p}+\frac{1}{q} \\ &= 1.\end{aligned}$$

即

$$\sum_{i=1}^{n}|x_i|\cdot|y_i| \leq \left(\sum_{i=1}^{n}|x_i|^p\right)^{1/p}\left(\sum_{i=1}^{n}|y_i|^q\right)^{1/q}.$$

例 2 设 $x=(x_1, x_2, \cdots, x_n)^T \in \mathbf{C}^n$，则

$$\|x\|_p=\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}, \quad 1 \leq p < \infty,$$

是 $\mathbf{C}^n$ 上的向量范数，称为 *Hölder* 范数。

证明：正定性和齐次性容易得证。当 $p=1$ 时为向量 1 范数，前面已经证明，下面只考虑 $p>1$ 情况。

三角不等式的证明。因为

$$\begin{aligned}\sum_{i=1}^{n}(|x_i|+|y_i|)^p &= \sum_{i=1}^{n}(|x_i|+|y_i|)(|x_i|+|y_i|)^{p-1} \\ &= \sum_{i=1}^{n}|x_i|(|x_i|+|y_i|)^{p-1}+\sum_{i=1}^{n}|y_i|(|x_i|+|y_i|)^{p-1}\end{aligned}$$

$$
\begin{aligned}
&\leq \left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}\left[\sum_{i=1}^{n}(|x_i|+|y_i|)^{(p-1)q}\right]^{\frac{1}{q}}+\left(\sum_{i=1}^{n}|y_i|^p\right)^{\frac{1}{p}}\left[\sum_{i=1}^{n}(|x_i|+|y_i|)^{(p-1)q}\right]^{\frac{1}{q}}\\
&=\left[\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}+\left(\sum_{i=1}^{n}|y_i|^p\right)^{\frac{1}{p}}\right]\left[\sum_{i=1}^{n}(|x_i|+|y_i|)^{(p-1)q}\right]^{\frac{1}{q}}\\
&=\left[\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}+\left(\sum_{i=1}^{n}|y_i|^p\right)^{\frac{1}{p}}\right]\left[\sum_{i=1}^{n}(|x_i|+|y_i|)^{p}\right]^{\frac{1}{q}}.\ (\text{因}(p-1)q=p.)
\end{aligned}
$$

即

$$\left(\sum_{i=1}^{n}(|x_i|+|y_i|)^p\right)^{\frac{1}{p}}\leq\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}+\left(\sum_{i=1}^{n}|y_i|^p\right)^{\frac{1}{p}}.$$

因为

$$\left(\sum_{i=1}^{n}(|x_i+y_i|)^p\right)^{\frac{1}{p}}\leq\left(\sum_{i=1}^{n}(|x_i|+|y_i|)^p\right)^{\frac{1}{p}},$$

故

$$\left(\sum_{i=1}^{n}(|x_i+y_i|)^p\right)^{\frac{1}{p}}\leq\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}+\left(\sum_{i=1}^{n}|y_i|^p\right)^{\frac{1}{p}},$$

即

$$\|x+y\|_p\leq\|x\|_p+\|y\|_p,$$

得证.

定理 2 设 $\|\cdot\|$ 是 $\mathbf{C}^m$ 上的范数，$A\in\mathbf{C}^{m\times n}$ 列满秩，则 $\|A\cdot\|$ 是 $\mathbf{C}^n$ 上的范数.

证明：

(1) 因 A 列满秩，则若 $x\neq\mathbf{0}$，有 $Ax\neq\mathbf{0}$，即 $\|Ax\|>0$.

(2) 对 $\forall\lambda\in\mathbf{C}$，则 $\|A(\lambda x)\|=\|\lambda Ax\|=|\lambda|\|Ax\|$.

(3) $\|A(x+y)\|=\|Ax+Ay\|\leq\|Ax\|+\|Ay\|$.

故满足范数定义.

定义 3 设 $x^{(k)}=\left(x_1^{(k)},x_2^{(k)},\cdots,x_n^{(k)}\right)^T\in\mathbf{C}^n$，如果

$$\lim_{k\to\infty}x_i^{(k)}=a_i,\quad(i=1,2,\cdots,n)$$

则向量序列 $x^{(k)}$ 收敛于 $a=(a_1,a_2,\cdots,a_n)^T$. 记为

$$\lim_{k\to\infty}x^{(k)}=a.$$

引理 2 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 为线性空间 $V_n(\mathbf{P})$ 的一组标准正交基,$x=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\,\tilde{x},\tilde{x}=(x_1,\cdots,x_n)^T\in\mathbf{P}^n$, 则 $V_n(\mathbf{P})$ 上的向量范数 $\|x\|$ 在闭球

$$S=\left\{x|(\tilde{x},\tilde{x})^{\frac{1}{2}}\leq 1\right\}$$

上有界.

证明: 因为 $x=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\,\widetilde{x}\in S$，故有

$$(\tilde{x},\tilde{x})=\sum_{i=1}^{n}|x_i|^2\leq 1,$$

即 $|x_i|\leq 1,\quad (i=1,2,\cdots,n)$.

记 $\sum_{i=1}^{n}\|\varepsilon_i\|=M$ 为正常数，则

$$\|x\|=\left\|\sum_{i=1}^{n}x_i\varepsilon_i\right\|\leq\sum_{i=1}^{n}\|x_i\varepsilon_i\|=\sum_{i=1}^{n}|x_i|\,\|\varepsilon_i\|\leq\sum_{i=1}^{n}\|\varepsilon_i\|=M,$$

故 $\|x\|$ 在 S 上有界.

引理 3 设 $\|x\|$ 是 $V_n(\mathbf{P})$ 上的向量范数，则 $\|x\|$ 是关于 $\|x\|_2$ 的连续函数.

证明: 设非零向量 $\Delta\tilde{x}=(\Delta\tilde{x}_1,\Delta\tilde{x}_2,\cdots,\Delta\tilde{x}_n)^T\in\mathbf{P}^n$,且 $\Delta x=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\,\Delta\tilde{x}\in V_n(\mathbf{P})$，则根据三角不等式有

$$|\|x+\Delta x\|-\|x\||\leq\|\Delta x\|=\|\Delta x\,\|_2\|\,\frac{\Delta x}{\|\Delta x\|_2}\|.$$

由于 $\frac{\Delta x}{\|\Delta x\|_2}\in S$，有引理 2 知 $\|\frac{\Delta x}{\|\Delta x\|_2}\|$ 有界.

又因为

$$\lim_{\Delta x\to 0}\|\Delta x\|_2=0,$$

所以

$$\lim_{\Delta x\to 0}|\|x+\Delta x\|-\|x\||=0,$$

故 $\|x\|$ 是关于 $\|x\|_2$ 的连续函数.

定义 4 设在 $V_n(\mathbf{P})$ 上定义了 $\|x\|_a\|$, $x\|_b$ 两种向量范数，若存在常数 $m>0$, $M>0$, 使得

$$m\|x\|_a\leq\|x\|_b\leq M\|x\|_a,\quad\forall x\in V_n(\mathbf{P}),$$

则称 $\|x\|_a$ 与 $\|x\|_b$ 等价.

等价关系： 自反性，对称性和传递性。

定理 3 $V_n(\mathbf{P})$ 上的任意两个向量范数均等价。

证明: 设 $\|x\|_a, \|x\|_b$ 为任意两向量函数，由引理 3 知 $\|x\|_a, \|x\|_b$ 都是 $\|x\|_2$ 的连续函数. 若 $x = 0$，显然满足等价的定义.

若 $x \neq 0$，定义函数

$$\phi(x) = \frac{\|x\|_a}{\|x\|_b} = \frac{\left\|\frac{x}{\|x\|_2}\right\|_a}{\left\|\frac{x}{\|x\|_2}\right\|_b},$$

易知 $\phi(x)$ 是单位球 S(有界闭集) 上关于 $\|x\|_2$ 的连续函数，故 $\exists M > 0$，使得 $\phi(x) = \frac{\|x\|_a}{\|x\|_b} \leq M$，即

$$\|x\|_a \leq M\|x\|_b.$$

同理，$\exists m > 0$ 可证

$$\|x\|_b \leq m\|x\|_a.$$

因此，$V_n(\mathbf{P})$ 上的任意两个向量范数均等价.

定理 4 设 $\|\cdot\|$ 是 $\mathbf{C}^n$ 上的任一向量范数，则 $\lim\limits_{k\to\infty} x^{(k)} = a$ 等价于 $\lim\limits_{k\to\infty} \left\|x^{(k)} - a\right\| = 0$.

证明: 因为

$$\begin{aligned}
\lim_{k\to\infty} x^{(k)} = a &\Longleftrightarrow \lim_{k\to\infty} x_i^{(k)} = a_i, \quad (1 \leq i \leq n) \\
&\Longleftrightarrow \lim_{k\to\infty} \left|x_i^{(k)} - a_i\right| = 0, \quad (1 \leq i \leq n) \\
&\Longleftrightarrow \lim_{k\to\infty} \max_{1\leq i\leq n} \left\{\left|x_i^{(k)} - a_i\right|\right\} = 0 \\
&\Longleftrightarrow \lim_{k\to\infty} \left\|x^{(k)} - a\right\|_\infty = 0.
\end{aligned}$$

再由 $V_n(\mathbf{P})$ 上的任意两个向量范数均等价和夹逼准则，有

$$\lim_{k\to\infty} \left\|x^{(k)} - a\right\| = 0,$$

得证.

例 3 令 $a = (1, 1, \cdots, 1)^T \in \mathbf{R}^n$，且

$$x^{(k)} = \left(1 + \frac{1}{2^k}, 1 + \frac{1}{3^k}, \cdots, 1 + \frac{1}{(n+1)^k}\right)^T$$

则

$$\lim_{k\to\infty} x^{(k)} = a.$$

证明：因为

$$\begin{aligned}\lim_{k\to\infty}\left\|x^{(k)}-a\right\|_\infty &= \lim_{k\to\infty}\max_{1\le i\le n}\left\{\left|x_i^{(k)}-a_i\right|\right\}\\ &= \lim_{k\to\infty}\max_{1\le i\le n}\left\{\frac{1}{(i+1)^k}\right\}\\ &= \lim_{k\to\infty}\frac{1}{2^k}\\ &= 0,\end{aligned}$$

故得证.

§2 矩阵的范数

程光辉

2020 年 3 月 28 日

定义 1 设 $\boldsymbol{A}=(\boldsymbol{a_{ij}})\in\mathbf{P}^{m\times n}$, 若映射 $\|\cdot\|:\mathbf{P}^{m\times n}\to\mathbf{R}$ 满足:

(1) 正定性 $\|\boldsymbol{A}\|\geq\mathbf{0}$，当且仅当 $\boldsymbol{A}=\mathbf{0}$ 时，$\|\boldsymbol{A}\|=\mathbf{0}$;

(2) 齐次性 $\|\boldsymbol{\lambda A}\|=|\boldsymbol{\lambda}|\|\boldsymbol{A}\|,\quad\forall\boldsymbol{\lambda}\in\mathbf{P},\ \forall\boldsymbol{A}\in\mathbf{P}^{m\times n}$;

(3) 三角不等式 $\|\boldsymbol{A}+\boldsymbol{B}\|\leq\|\boldsymbol{A}\|+\|\boldsymbol{B}\|,\forall\boldsymbol{A},\boldsymbol{B}\in\mathbf{P}^{m\times n}$,

则称映射 $\|\cdot\|$ 为 $\mathbf{P}^{m\times n}$ 上的矩阵范数.

例 1 设 $\boldsymbol{A}\in\mathbf{P}^{m\times n}$，则

$$\|\boldsymbol{A}\|_{m_1}=\sum_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|,$$

$$\|\boldsymbol{A}\|_{m_2}=\left(\sum_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|^2\right)^{\frac{1}{2}},$$

$$\|\boldsymbol{A}\|_{m_\infty}=\max_{i,j}\{|a_{ij}|\},\quad 1\leq i\leq m,\quad 1\leq j\leq n,$$

都是矩阵范数.

显然，$\|\boldsymbol{A}\|_{m_1}$，$\|\boldsymbol{A}\|_{m_2}$，$\|\boldsymbol{A}\|_{m_\infty}$ 是向量范数 $\|\cdot\|_1$，$\|\cdot\|_2$，$\|\cdot\|_\infty$ 的自然推广.

定理 1 $\mathbf{P}^{m\times n}$ 上的任意两个矩阵范数均等价。

定义 2 设 $\|\cdot\|_a:\mathbf{P}^{m\times l}\to\mathbf{R},\quad\|\cdot\|_b:\mathbf{P}^{l\times n}\to\mathbf{R},\quad\|\cdot\|_c:\mathbf{P}^{m\times n}\to\mathbf{R}$ 是矩阵范数，如果

$$\|\boldsymbol{AB}\|_c\leq\|\boldsymbol{A}\|_a\cdot\|\boldsymbol{B}\|_b,$$

则称矩阵范数 $\|\cdot\|_a,\quad\|\cdot\|_b$ 和 $\|\cdot\|_c$ 相容. 如果

$$\|\boldsymbol{AB}\|\leq\|\boldsymbol{A}\|\cdot\|\boldsymbol{B}\|,$$

则称 $\|\cdot\|$ 是自相容矩阵范数.

例 2 设 $A \in \mathbf{P}^{n\times n}$，矩阵范数 $\|A\|_{m_\infty} = \max\limits_{i,j}\{|a_{ij}|\}$， $(1 \le i,\ j \le n)$ 是不相容的矩阵范数.

解： 取

$$A = B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

则

$$AB = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix},$$

即有

$$\|AB\|_{m_\infty} = 2 \nleq \|A\|_{m_\infty}\|B\|_{m_\infty} = 1,$$

因此，矩阵范数 $\|\cdot\|_{m_\infty}$ 是不相容的矩阵范数.

例 3 $\|\cdot\|_{m_1}$ 和 $\|\cdot\|_{m_2}$ 是相容的矩阵范数.

证明： 设 $A \in \mathbf{P}^{m\times l}$， $B \in \mathbf{P}^{l\times n}$，则

$$\begin{aligned}
\|AB\|_{m_1} &= \sum_{j=1}^{n}\sum_{i=1}^{m}\left|\sum_{k=1}^{l} a_{ik}b_{kj}\right| \\
&\le \sum_{j=1}^{n}\sum_{i=1}^{m}\sum_{k=1}^{l} |a_{ik}|\cdot|b_{kj}| \\
&= \sum_{k=1}^{l}\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ik}|\cdot|b_{kj}| \\
&= \sum_{k=1}^{l}\left(\sum_{i=1}^{m}|a_{ik}|\cdot\sum_{j=1}^{n}|b_{kj}|\right) \\
&\le \sum_{k=1}^{l}\left(\sum_{i=1}^{m}\sum_{k=1}^{l}|a_{ik}|\cdot\sum_{j=1}^{n}|b_{kj}|\right) \\
&= \left(\sum_{i=1}^{m}\sum_{k=1}^{l}|a_{ik}|\right)\cdot\left(\sum_{k=1}^{l}\sum_{j=1}^{n}|b_{kj}|\right) \\
&= \|A\|_{m_1}\cdot\|B\|_{m_1}.
\end{aligned}$$

$$\|AB\|_{m_2} = \left(\sum_{j=1}^{n}\sum_{i=1}^{m}\left|\sum_{k=1}^{l} a_{ik}b_{kj}\right|^2\right)^{\frac{1}{2}}$$

$$
\begin{aligned}
&\le \left[\sum_{j=1}^{n}\sum_{i=1}^{m}\left(\sum_{k=1}^{l}|a_{ik}|\cdot|b_{kj}|\right)^{2}\right]^{\frac{1}{2}}\\
&\le \left[\sum_{j=1}^{n}\sum_{i=1}^{m}\left(\sum_{k=1}^{l}|a_{ik}|^{2}\right)\cdot\left(\sum_{k=1}^{l}|b_{kj}|^{2}\right)\right]^{\frac{1}{2}}\\
&= \left\{\sum_{j=1}^{n}\left(\sum_{k=1}^{l}|b_{kj}|^{2}\right)\cdot\left[\sum_{i=1}^{m}\left(\sum_{k=1}^{l}|a_{ik}|^{2}\right)\right]\right\}^{\frac{1}{2}}\\
&= \left(\sum_{i=1}^{m}\sum_{k=1}^{l}|a_{ik}|^{2}\right)^{\frac{1}{2}}\cdot\left(\sum_{k=1}^{l}\sum_{j=1}^{n}|b_{kj}|^{2}\right)^{\frac{1}{2}}\\
&= \|A\|_{m_2}\cdot\|B\|_{m_2}.
\end{aligned}
$$

得证.

例 4 设 $A\in \mathbf{P}^{n\times n}$，则

$$
\|A\|_a = n\max_{i,j}\{|a_{ij}|\},\quad 1\le i,j\le n,
$$

是相容的矩阵范数。

证明： $\|\cdot\|_a$ 是矩阵范数很显然.

下面证明相容性，令 $A,B\in \mathbf{P}^{n\times n}$，

$$
\begin{aligned}
\|AB\|_a &= n\max_{i,j}\left\{\left|\sum_{k=1}^{n}a_{ik}b_{kj}\right|\right\}\\
&\le n\max_{i,j}\left\{\sum_{k=1}^{n}|a_{ik}||b_{kj}|\right\}\\
&\le n\max_{i,j}\left\{n\max_{k}|a_{ik}||b_{kj}|\right\}\\
&\le n\max_{i,k}\{|a_{ik}|\}\cdot n\max_{k,j}\{|b_{kj}|\}\\
&= \|A\|_a\cdot\|B\|_a,
\end{aligned}
$$

得证.

定理 2 设 $A\in \mathbf{P}^{n\times n}$，

(1) 若 $A=(a_1,a_2,\cdots,a_n)$，则

$$
\|A\|_F^2=\|A\|_{m_2}^2=\sum_{i=1}^{n}\|a_i\|_2^2
$$

其中 $\|a_i\|_2^2=a_i^H a_i$.

(2) $\|\boldsymbol{A}\|_{m_2}^2 = \mathbf{tr}\left(\boldsymbol{A}^H\boldsymbol{A}\right) = \sum\limits_{i=1}^{n} \boldsymbol{\lambda}_i\left(\boldsymbol{A}^H\boldsymbol{A}\right)$.

(3) 对任意的酉矩阵 $\boldsymbol{U},\boldsymbol{V}\in\mathbf{C}^{n\times n}$，有 $\|\boldsymbol{A}\|_{m_2}^2 = \left\|\boldsymbol{U}^H\boldsymbol{A}\boldsymbol{V}\right\|_{m_2}^2 = \left\|\boldsymbol{U}\boldsymbol{A}\boldsymbol{V}^H\right\|_{m_2}^2$.

证明： (1)，(2) 比较显然.

下面只对 (3) 进行证明.

$$\begin{aligned}
\|\boldsymbol{A}\|_{m_2}^2 &= \mathbf{tr}\left(\boldsymbol{A}^H\boldsymbol{A}\right)\\
&= \mathbf{tr}\left(\boldsymbol{A}\boldsymbol{A}^H\right)\\
&= \mathbf{tr}\left(\boldsymbol{A}\boldsymbol{V}\boldsymbol{V}^H\boldsymbol{A}^H\right)\\
&= \mathbf{tr}\left[\boldsymbol{A}\boldsymbol{V}(\boldsymbol{A}\boldsymbol{V})^H\right]\\
&= \mathbf{tr}\left[(\boldsymbol{A}\boldsymbol{V})^H\boldsymbol{A}\boldsymbol{V}\right]\\
&= \mathbf{tr}\left(\boldsymbol{V}^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{V}\right)\\
&= \mathbf{tr}\left(\boldsymbol{V}^H\boldsymbol{A}^H\boldsymbol{U}\boldsymbol{U}^H\boldsymbol{A}\boldsymbol{V}\right)\\
&= \mathbf{tr}\left[\left(\boldsymbol{U}^H\boldsymbol{A}\boldsymbol{V}\right)^H\left(\boldsymbol{U}^H\boldsymbol{A}\boldsymbol{V}\right)\right]\\
&= \left\|\boldsymbol{U}^H\boldsymbol{A}\boldsymbol{V}\right\|_{m_2}^2,
\end{aligned}$$

得证.

推论 1 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，对任意的酉矩阵 $\boldsymbol{U},\boldsymbol{V}\in\mathbf{C}^{n\times n}$，有

$$\|\boldsymbol{A}\|_{m_2} = \|\boldsymbol{U}\boldsymbol{A}\|_{m_2} = \|\boldsymbol{A}\boldsymbol{V}\|_{m_2} = \|\boldsymbol{U}\boldsymbol{A}\boldsymbol{V}\|_{m_2}.$$

§3 算子范数

程光辉

2020 年 3 月 28 日

定义 1 设 $\|\bullet\|_a$ 是 $\mathbf{P}^n$ 上的向量范数，$\|*\|_m$ 是 $\mathbf{P}^{n\times n}$ 上的矩阵范数，且

$$\|\boldsymbol{A}\boldsymbol{x}\|_a \leq \|\boldsymbol{A}\|_m \|\boldsymbol{x}\|_a,$$

则称 $\|*\|_m$ 为与向量范数 $\|\bullet\|_a$ 相容的矩阵范数.

例 1 设 $\boldsymbol{x}\in\mathbf{P}^n$，$\boldsymbol{A}=(a_{ij})\in\mathbf{P}^{n\times n}$，则

$$\|\boldsymbol{A}\|_{m_1}=\sum_{j=1}^{n}\sum_{i=1}^{n}|a_{ij}|$$

是与向量范数 $\|\bullet\|_1$ 相容的矩阵范数.

证明： 因为

$$\begin{aligned}
\|\boldsymbol{A}\boldsymbol{x}\|_1 &= \sum_{i=1}^{n}\left|\sum_{k=1}^{n}a_{ik}x_k\right| \\
&\leq \sum_{i=1}^{n}\sum_{k=1}^{n}|a_{ik}x_k| \\
&= \sum_{k=1}^{n}\sum_{i=1}^{n}|a_{ik}|\,|x_k| \\
&= \sum_{k=1}^{n}\left(|x_k|\cdot\sum_{i=1}^{n}|a_{ik}|\right) \\
&\leq \sum_{k=1}^{n}\left(\sum_{k=1}^{n}|x_k|\cdot\sum_{i=1}^{n}|a_{ik}|\right) \\
&= \left(\sum_{i=1}^{n}\sum_{k=1}^{n}|a_{ik}|\right)\cdot\sum_{k=1}^{n}|x_k| \\
&= \|\boldsymbol{A}\|_{m_1}\bullet\|\boldsymbol{x}\|_1,
\end{aligned}$$

所以，$\|\boldsymbol{A}\|_{m_1}$ 是与向量范数 $\|\bullet\|_1$ 相容的矩阵范数.

例 2 设 $x \in \mathbf{P}^n, A = (a_{ij}) \in \mathbf{P}^{n\times n}$，则 $\|A\|_{m_2}$ 是与 $\|x\|_2$ 相容的矩阵范数.

证明： 因为

$$
\begin{aligned}
\|Ax\|_2^2 &= \sum_{i=1}^{n} |a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n|^2 \\
&\le \sum_{i=1}^{n} \left(\sum_{j=1}^{n} |a_{ij}|^2\right) \cdot \left(\sum_{j=1}^{n} |x_j|^2\right) \\
&= \left(\sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^2\right) \cdot \sum_{j=1}^{n} |x_j|^2 \\
&= \|A\|_{m_2}^2 \cdot \|x\|_2^2,
\end{aligned}
$$

所以，$\|A\|_{m_2}$ 是与 $\|x\|_2$ 相容的矩阵范数.

对任意向量范数，是否存在与之相容的矩阵范数？

定理 1 设 $\|x\|_a$ 是 $\mathbf{P}^n$ 上的向量范数，$A \in \mathbf{P}^{n\times n}$，则

$$
\|A\|_a = \max_{x\neq 0} \frac{\|Ax\|_a}{\|x\|_a} \left(= \max_{\|u\|_a=1} \|Au\|_a\right)
$$

是与向量范数 $\|x\|_a$ 相容的矩阵范数。

证明：

(1) 正定性: 因为 $A \neq 0$，则 $\mathbf{P}^n$ 中存在 $x_0 \neq 0$，使 $Ax_0 \neq 0$，那么有 $\|Ax_0\|_a > 0, \|x_0\|_a > 0$，则

$$
\|A\|_a \ge \frac{\|Ax_0\|_a}{\|x_0\|_a} > 0.
$$

(2) 齐次性:

$$
\begin{aligned}
\|\lambda A\|_a &= \max_{x\neq 0} \frac{\|\lambda Ax\|_a}{\|x\|_a} \\
&= \max_{x\neq 0} \frac{|\lambda| \cdot \|Ax\|_a}{\|x\|_a} \\
&= |\lambda| \cdot \max_{x\neq 0} \frac{\|Ax\|_a}{\|x\|_a} \\
&= |\lambda| \cdot \|A\|_a.
\end{aligned}
$$

(3) 三角不等式:

$$
\|A+B\|_a = \max_{x\neq 0} \frac{\|(A+B)x\|_a}{\|x\|_a}
$$

$$
\begin{aligned}
&\le \max_{x\neq 0}\frac{\|\boldsymbol{Ax}\|_a+\|\boldsymbol{Bx}\|_a}{\|\boldsymbol{x}\|_a}\\
&\le \max_{x\neq 0}\frac{\|\boldsymbol{Ax}\|_a}{\|\boldsymbol{x}\|_a}+\max_{x\neq 0}\frac{\|\boldsymbol{Bx}\|_a}{\|\boldsymbol{x}\|_a}\\
&=\|\boldsymbol{A}\|_a+\|\boldsymbol{B}\|_a.
\end{aligned}
$$

相容性:
因为

$$\|\boldsymbol{A}\|_a=\max_{x\neq 0}\frac{\|\boldsymbol{Ax}\|_a}{\|\boldsymbol{x}\|_a},$$

则

$$\|\boldsymbol{A}\|_a\ge \frac{\|\boldsymbol{Ax}\|_a}{\|\boldsymbol{x}\|_a},$$

故

$$\|\boldsymbol{Ax}\|_a\le \|\boldsymbol{A}\|_a\cdot\|\boldsymbol{x}\|_a.$$

推论 1 设 $\|\boldsymbol{x}\|_a$ 是 $\mathbf{P}^n$ 上的向量范数，$\boldsymbol{A},\boldsymbol{B}\in\mathbf{P}^{n\times n}$，$\|\boldsymbol{A}\|_a$ 是从属于 $\|\boldsymbol{x}\|_a$ 的算子范数，则它是相容的矩阵范数，即

$$\|\boldsymbol{AB}\|_a\le \|\boldsymbol{A}\|_a\cdot\|\boldsymbol{B}\|_a.$$

证明：因为

$$
\begin{aligned}
\|\boldsymbol{AB}\|_a&=\max_{x\neq 0}\frac{\|\boldsymbol{ABx}\|_a}{\|\boldsymbol{x}\|_a}\\
&\le \max_{x\neq 0}\frac{\|\boldsymbol{A}\|_a\cdot\|\boldsymbol{Bx}\|_a}{\|\boldsymbol{x}\|_a}\\
&=\|\boldsymbol{A}\|_a\cdot\max_{x\neq 0}\frac{\|\boldsymbol{Bx}\|_a}{\|\boldsymbol{x}\|_a}\\
&=\|\boldsymbol{A}\|_a\cdot\|\boldsymbol{B}\|_a,
\end{aligned}
$$

得证.

算子范数的特性

(1) 它是所有与向量范数 $\|\boldsymbol{x}\|_a$ 相容的矩阵范数中最小的.

证明：因为

$$\|\boldsymbol{Ax}\|_a\le \|\boldsymbol{A}\|\cdot\|\boldsymbol{x}\|_a\quad(\forall \boldsymbol{x}\in\mathbf{P}^n),$$

则

$$\frac{\|\boldsymbol{Ax}\|_a}{\|\boldsymbol{x}\|_a}\le \|\boldsymbol{A}\|\quad(\forall \boldsymbol{0}\neq \boldsymbol{x}\in\mathbf{P}^n),$$

故

$$\|\boldsymbol{A}\|_a=\max_{x\neq 0}\frac{\|\boldsymbol{Ax}\|_a}{\|\boldsymbol{x}\|_a}\le \|\boldsymbol{A}\|.$$

(2) 它的两种表达形式

$$\|A\|_a = \max_{x\neq 0} \frac{\|Ax\|_a}{\|x\|_a} \left(= \max_{\|u\|_a=1} \|Au\|_a \right).$$

(3) 它是自相容矩阵范数 (由推论 1).

定理 2 设 $\|\bullet\|_m$ 是相容的矩阵范数，则存在向量范数 $\|x\|$，使

$$\|Ax\| \leq \|A\|_m \cdot \|x\|.$$

证明：定义如下映射，证明此映射为向量范数即可.

$$\|x\| = \left\|xa^H\right\|_m, \quad 0 \neq a \in \mathbf{P}^n, \forall x \in \mathbf{P}^n$$

(1) 因为 $a \neq 0$，对任意非零向量 $x \in \mathbf{P}^n$，则有 $xa^H \neq 0$，即 $\|x\| = \left\|xa^H\right\|_m > 0$.

(2) $\|\lambda x\| = \left\|\lambda x a^H\right\|_m = |\lambda| \cdot \left\|xa^H\right\|_m$.

(3) $\|x+y\| = \left\|(x+y)a^H\right\|_m = \left\|xa^H + ya^H\right\|_m \leq \left\|xa^H\right\|_m + \left\|ya^H\right\|_m = \|x\| + \|y\|$.

(4) $\|Ax\| = \left\|Axa^H\right\|_m \leq \|A\|_m \cdot \left\|xa^H\right\|_m = \|A\|_m \cdot \|x\|$.

例 3 取 $a = (1, 0, \cdots, 0)^T, x = (x_1, x_2, \cdots, x_n)^T$，则

$$\|x\| = \left\|xa^H\right\|_{m_2} = \left\| \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ x_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ x_n & 0 & \cdots & 0 \end{bmatrix} \right\|_{m_2} = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2} = \|x\|_2.$$

定理 3 如果 $\|\bullet\|_m : \mathbf{C}^{n\times n} \to \mathbf{R}$ 是一相容的矩阵范数，则对任意 $A \in \mathbf{C}^{n\times n}$，有

$$|\lambda| \leq \|A\|_m,$$

其中 λ 是 A 的特征值.

证明：因为 $Ax = \lambda x$，则

$$|\lambda| \cdot \|x\| = \|\lambda x\| = \|Ax\| \leq \|A\|_m \|x\|,$$

故

$$|\lambda| \leq \|A\|_m.$$

例 4 设 $A \in \mathbf{C}^{m\times n}$，则从属于向量范数 $\|x\|_1$ 的算子范数为

$$\|A\|_1 = \max_{1\leq j\leq n} \left\{ \sum_{i=1}^{n} |a_{ij}| \right\},$$

称为极大列和范数.

证明：设 $\boldsymbol{A}=[\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n]$，则

$$
\begin{aligned}
\|\boldsymbol{Ax}\|_1 &= \|x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\cdots+x_n\boldsymbol{a}_n\|_1\\
&\le \sum_{i=1}^{n}\|x_i\boldsymbol{a}_i\|_1\\
&= \sum_{i=1}^{n}|x_i|\cdot\|\boldsymbol{a}_i\|_1\\
&\le \sum_{i=1}^{n}|x_i|\left(\max_{1\le i\le n}\{\|\boldsymbol{a}_i\|_1\}\right)\\
&= \max_{1\le i\le n}\{\|\boldsymbol{a}_i\|_1\}\,\|\boldsymbol{x}\|_1,
\end{aligned}
$$

既有

$$
\|\boldsymbol{A}\|_1=\max_{x\neq 0}\frac{\|\boldsymbol{Ax}\|_1}{\|\boldsymbol{x}\|_1}\le\max_{1\le i\le n}\{\|\boldsymbol{a}_i\|_1\}, \tag{1}
$$

不妨令

$$
\|\boldsymbol{a}_k\|_1=\max_{1\le i\le n}\{\|\boldsymbol{a}_i\|_1\}=\max_{1\le i\le n}\left\{\sum_{i=1}^{n}|a_{ij}|\right\},
$$

若取 $\tilde{\boldsymbol{x}}=\begin{bmatrix}0 & \cdots & 0 & \lambda & 0 & \cdots & 0\end{bmatrix}^T$ ($\boldsymbol{\lambda}$ 位于第 $\boldsymbol{k}$ 个位置), 则

$$
\max_{x\neq 0}\|\boldsymbol{Ax}\|_1\ge\|\boldsymbol{A}\tilde{\boldsymbol{x}}\|_1=\|\boldsymbol{\lambda a}_k\|_1=\|\boldsymbol{a}_k\|_1\cdot\|\tilde{\boldsymbol{x}}\|_1,
$$

因此，

$$
\|\boldsymbol{A}\|_1=\max_{x\neq 0}\frac{\|\boldsymbol{Ax}\|_1}{\|\boldsymbol{x}\|_1}\ge\frac{\|\boldsymbol{A}\tilde{\boldsymbol{x}}\|_1}{\|\tilde{\boldsymbol{x}}\|_1}=\|\boldsymbol{a}_k\|_1=\max_{1\le i\le n}\left\{\sum_{i=1}^{n}|a_{ij}|\right\}. \tag{2}
$$

所以，由 **(1), (2)** 得证.

例 5 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{m\times n}$，则从属于 $\|\boldsymbol{x}\|_\infty$ 的算子范数

$$
\|\boldsymbol{A}\|_\infty=\max_{1\le i\le m}\left(\sum_{j=1}^{n}|a_{ij}|\right),
$$

称为极大行和范数.

证明：令 $\boldsymbol{x}\in\mathbf{C}^n$，因为

$$
\begin{aligned}
\|\boldsymbol{Ax}\|_\infty &= \max_{1\le i\le m}\left\{\left|\sum_{j=1}^{n}a_{ij}x_j\right|\right\}\\
&\le \max_{1\le i\le m}\left\{\sum_{j=1}^{n}|a_{ij}|\cdot|x_j|\right\}
\end{aligned}
$$

$$\le \max_{1\le i\le m}\left\{\sum_{j=1}^{n}|a_{ij}|\cdot \max_{1\le j\le n}|x_j|\right\}$$
$$= \max_{1\le i\le m}\left\{\sum_{j=1}^{n}|a_{ij}|\right\}\cdot \max_{1\le j\le n}\{|x_j|\}$$

则
$$\|A\|_\infty=\max_{x\neq 0}\frac{\|Ax\|_\infty}{\|x\|_\infty}\le \max_{1\le i\le m}\left\{\sum_{j=1}^{n}|a_{ij}|\right\}. \tag{3}$$

不妨令
$$\sum_{j=1}^{n}|a_{kj}|=\max_{1\le i\le m}\left(\sum_{j=1}^{n}|a_{ij}|\right).$$

若 $A=0$，显然.

假设 $A\neq 0$, $\lambda>0$, 令 $z=(z_1,z_2,\cdots,z_n)^T\in \mathrm{C}^n$，其中
$$\begin{cases} z_j=\frac{\lambda|a_{kj}|}{a_{kj}}, & (a_{kj}\neq 0)\\ z_j=\lambda, & (a_{kj}=0)\end{cases}$$

则 $\|z\|_\infty=\lambda,\quad a_{kj}z_j=\lambda\,|a_{kj}|\,,\quad (j=1,\cdots,n)$.

因为
$$\begin{aligned}\max_{x\neq 0}\|Ax\|_\infty &\ge \|Az\|_\infty\\ &=\max_{1\le i\le m}\left|\sum_{j=1}^{n}a_{ij}z_j\right|\\ &\ge \left|\sum_{j=1}^{n}a_{kj}z_j\right|\\ &=\lambda\sum_{j=1}^{n}|a_{kj}|\\ &=\|z\|_\infty\cdot\sum_{j=1}^{n}|a_{kj}|\,,\end{aligned}$$

所以
$$\|A\|_\infty=\max_{x\neq 0}\frac{\|Ax\|_\infty}{\|x\|_\infty}\ge\frac{\|Az\|_\infty}{\|z\|_\infty}\ge\sum_{j=1}^{n}|a_{kj}|=\max_{1\le i\le m}\left\{\sum_{j=1}^{n}|a_{ij}|\right\}. \tag{4}$$

综上，由 (3), (4) 得证.

定义 2 设 $A\in\mathrm{C}^{n\times n}$, λ_i 是 A 的特征值，则 $r(A)=\max\limits_{1\le i\le n}\{|\lambda_i|\}$ 称为 A 的谱半径.

例 6 设 $A \in \mathbf{P}^{m\times n}$，则从属于 $\|x\|_2$ 的算子范数 *(又称为谱范数)* 为

$$\|A\|_2 = \sqrt{r\left(A^H A\right)}.$$

证明： 因为对任意非零向量 $X \in \mathbf{C}^n$，都有

$$f(X) = X^H\left(A^H A\right)X = (AX)^H AX \geq 0,$$

则 $A^H A$ 为半正定 Hermitian 矩阵，特征值非负.

令矩阵 $A^H A$ 的特征值为 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$，$X_i$ 是对应 λ_i 的单位正交特征向量，对 $\forall u \in \mathbf{P}^n$ 且 $\|u\|_2 = 1$，有

$$u = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n,$$

那么

$$\|u\|_2^2 = u^H u = |a_1|^2 + |a_2|^2 + \cdots + |a_n|^2 = 1.$$

又

$$A^H A u = a_1\lambda_1 X_1 + a_2\lambda_2 X_2 + \cdots + a_n\lambda_n X_n,$$

即

$$\begin{aligned}
\|Au\|_2^2 &= u^H A^H A u \\
&= \lambda_1 |a_1|^2 + \lambda_2 |a_2|^2 + \cdots + \lambda_n |a_n|^2 \\
&\leq \lambda_1\left(|a_1|^2 + |a_2|^2 + \cdots + |a_n|^2\right) \\
&= \lambda_1.
\end{aligned}$$

因此

$$\max_{\|u\|_2=1} \|Au\|_2 \leq \sqrt{\lambda_1}.$$

又因为

$$\|AX_1\|_2^2 = X_1^H A^H A X_1 = X_1^H \lambda_1 X_1 = \lambda_1,$$

故

$$\|A\|_2 = \max_{\|u\|_2=1} \|Au\|_2 = \sqrt{\lambda_1} = \sqrt{r\left(A^H A\right)}.$$

谱范数的性质

定理 4 设 $A \in \mathbf{C}^{n\times n}$, 则

(1) $\|A\|_2 = \left\|A^H\right\|_2 = \left\|A^T\right\|_2 = \|\overline{A}\|_2$.

(2) $\left\|A^H A\right\|_2 = \left\|A A^H\right\|_2 = \|A\|_2^2$.

(3) 对任意 n 阶酉矩阵 U 及 V 都有

$$\|UA\|_2 = \|AV\|_2 = \|UAV\|_2 = \|A\|_2.$$

证明：(1) 由 $A^HAx = \lambda x$，若 $\lambda = 0$，则 A^HA 非满秩，那么 AA^H 也非满秩，即 $\lambda = 0$ 也是 AA^H 的特征值.

若 $\lambda \neq 0$，则 $y = Ax \neq 0$，那么

$$AA^Hy = AA^HAx = A(\lambda x) = \lambda Ax = \lambda y,$$

则 λ 也是 AA^H 的特征值.

同理可证： AA^H 的特征值也是 A^HA 的特征值.

$$\|A\|_2 = \sqrt{r\left(A^HA\right)} = \sqrt{r\left(AA^H\right)} = \left\|A^H\right\|_2.$$

又因为

$$\begin{aligned}\left|\lambda E - \left(A^T\right)^H A^T\right| &= \left|\lambda E - \left(AA^H\right)^T\right| \\ &= \left|\lambda E - AA^H\right|,\end{aligned}$$

即 $\left(A^T\right)^H A^T$ 和 AA^H 的特征值相同.

故

$$\|A\|_2 = \left\|A^H\right\|_2 = \left\|A^T\right\|_2 = \|\overline{A}\|_2.$$

(2) 因为

$$\left\|A^HA\right\|_2^2 = r\left[\left(A^HA\right)^H\left(A^HA\right)\right] = r\left[\left(A^HA\right)^2\right] = \left[r\left(A^HA\right)\right]^2,$$

所以

$$\left\|A^HA\right\|_2 = \left\|AA^H\right\|_2 = \|A\|_2^2.$$

(3) $\|UA\|_2^2 = r\left[(UA)^H(UA)\right] = r\left[A^HU^HUA\right] = r\left(A^HA\right).$

定理 5 设 $A \in \mathbf{C}^{n\times n}$，则

(1) $\|A\|_2 = \max\limits_{\|x\|_2=\|y\|_2=1}\left|y^HAx\right|,$

(2) $\|A\|_2^2 \leq \|A\|_1\|A\|_\infty.$

证明：(1) 因为

$$\left|y^HAx\right| \leq \|y\|_2\|Ax\|_2 \leq \|y\|_2\|A\|_2\|x\|_2 = \|A\|_2,$$

所以，
$$\max_{\|x\|_2=\|y\|_2=1}\left|y^H A x\right| \leq \|A\|_2.$$
又因为
$$\|A\|_2=\max_{\|x\|_2=1}\|Ax\|_2,$$
则存在单位向量 x_0，使得
$$\|A\|_2=\|Ax_0\|_2>0.$$
取
$$y_0=\frac{Ax_0}{\|Ax_0\|_2},$$
因此
$$\left|y_0^H A x_0\right|=\left|\frac{(Ax_0)^H}{\|Ax_0\|_2}Ax_0\right|=\|Ax_0\|_2=\|A\|_2,$$
故
$$\max_{\|x\|_2=\|y\|_2=1}\left|y^H A x\right| = \|A\|_2.$$
(2) $\|A\|_2^2=r\left(A^HA\right)\leq\left\|A^HA\right\|_1\leq\left\|A^H\right\|_1\|A\|_1=\|A\|_1\|A\|_\infty.$

广义算子范数

定理 6 设 $\|\cdot\|_a,\|\cdot\|_b$ 都是向量范数，$A\in\mathbf{P}^{n\times n}$，则
$$\|A\|_{a,b}=\max_{x\neq 0}\frac{\|Ax\|_a}{\|x\|_b}\quad\left(=\max_{\|u\|_b=1}\|Au\|_a\right)$$
叫做 $\mathbf{P}^{n\times n}$ 上的广义算子范数.

定理 7 设 $\|\cdot\|_a,\|\cdot\|_b$ 与 $\|\cdot\|_c$ 都是向量范数，则
$$\|AB\|_{a,c}\leq\|A\|_{a,b}\|B\|_{b,c}.$$

§6 范数的应用

程光辉

2020 年 4 月 11 日

1　范数的应用

(1) 矩阵 $\boldsymbol{A}$ 可逆，$\boldsymbol{A}$ 与其扰动矩阵 $\boldsymbol{\delta A}$ 满足什么条件时，$\boldsymbol{A+\delta A}$ 可逆？
(2) 当 $\boldsymbol{A+\delta A}$ 可逆，$\boldsymbol{A^{-1}}$ 与 $\boldsymbol{(A+\delta A)^{-1}}$ 的近似程度如何估计？

例 1. 设

$$\boldsymbol{A=\begin{pmatrix} 2 & 6 \\ 2 & 6.00001 \end{pmatrix},\quad \delta A=\begin{pmatrix} 0 & 0 \\ 0 & -0.00002 \end{pmatrix}},$$

求 $\boldsymbol{A^{-1}}$，　$\boldsymbol{(A+\delta A)^{-1}}$.

解： 计算得

$$\boldsymbol{A^{-1}=\begin{pmatrix} 300000.5 & -300000 \\ -100000 & 100000 \end{pmatrix}},$$

$$\boldsymbol{(A+\delta A)^{-1}=\begin{pmatrix} -299999.5 & -300000 \\ 100000 & -100000 \end{pmatrix}}.$$

定义 1. 设 $\boldsymbol{A}$ 是可逆矩阵，称

$$\boldsymbol{K_p(A)=\|A\|_p\left\|A^{-1}\right\|_p}$$

是矩阵 $\boldsymbol{A}$ 的条件数.

定理 1. 设 $\boldsymbol{A\in \mathbf{C}^{n\times n}}$，$\boldsymbol{\|A\|_a}$ 是从属于向量范数 $\boldsymbol{\|x\|_a}$ 的算子范数，则当 $\boldsymbol{\|A\|_a<1}$ 时，$\boldsymbol{E-A}$ 可逆，且

$$\boldsymbol{\left\|(E-A)^{-1}\right\|_a\le(1-\|A\|_a)^{-1}}.$$

证明：对任意的非零向量 $\boldsymbol{x} \in \mathbf{C}^n$，有

$$\begin{aligned}\|(\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}\|_a &= \|\boldsymbol{x}-\boldsymbol{A}\boldsymbol{x}\|_a \\ &\geq \|\boldsymbol{x}\|_a - \|\boldsymbol{A}\boldsymbol{x}\|_a \\ &\geq \|\boldsymbol{x}\|_a - \|\boldsymbol{A}\|_a\|\boldsymbol{x}\|_a \\ &= \|\boldsymbol{x}\|_a\left(1-\|\boldsymbol{A}\|_a\right) \\ &> 0,\end{aligned}$$

即 $(\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x} \neq \mathbf{0}$，所以 $(\boldsymbol{E}-\boldsymbol{A})$ 可逆.

又因为

$$(\boldsymbol{E}-\boldsymbol{A})^{-1}(\boldsymbol{E}-\boldsymbol{A}) = \boldsymbol{E},$$

则

$$(\boldsymbol{E}-\boldsymbol{A})^{-1} = \boldsymbol{E} + \boldsymbol{A}(\boldsymbol{E}-\boldsymbol{A})^{-1}.$$

进而

$$\begin{aligned}\left\|(\boldsymbol{E}-\boldsymbol{A})^{-1}\right\|_a &= \left\|\boldsymbol{E}+\boldsymbol{A}(\boldsymbol{E}-\boldsymbol{A})^{-1}\right\|_a \\ &\leq \|\boldsymbol{E}\|_a + \|\boldsymbol{A}\|_a\left\|(\boldsymbol{E}-\boldsymbol{A})^{-1}\right\|_a,\end{aligned}$$

故

$$\left(1-\|\boldsymbol{A}\|_a\right)\left\|(\boldsymbol{E}-\boldsymbol{A})^{-1}\right\|_a \leq \|\boldsymbol{E}\|_a = 1.$$

整理即

$$\left\|(\boldsymbol{E}-\boldsymbol{A})^{-1}\right\|_a \leq \left(1-\|\boldsymbol{A}\|_a\right)^{-1},$$

得证.

定理 2. 设 $\boldsymbol{A}$ 是可逆,$\delta\boldsymbol{A}$ 为扰动矩阵, 且 $\left\|\boldsymbol{A}^{-1}\right\|_a\|\delta\boldsymbol{A}\|_a < 1$, 则

(1) $\boldsymbol{A}+\delta\boldsymbol{A}$ 可逆;

(2) $\frac{\left\|\boldsymbol{A}^{-1}-(\boldsymbol{A}+\delta\boldsymbol{A})^{-1}\right\|_a}{\left\|\boldsymbol{A}^{-1}\right\|_a} \leq \frac{K(\boldsymbol{A})\frac{\|\delta\boldsymbol{A}\|_a}{\|\boldsymbol{A}\|_a}}{1-K(\boldsymbol{A})\frac{\|\delta\boldsymbol{A}\|_a}{\|\boldsymbol{A}\|_a}}$.

证明：

(1) 由 $\left\|\boldsymbol{A}^{-1}\delta\boldsymbol{A}\right\|_a \leq \left\|\boldsymbol{A}^{-1}\right\|_a\|\delta\boldsymbol{A}\|_a < 1$，可得 $\boldsymbol{E}+\boldsymbol{A}^{-1}\delta\boldsymbol{A}$ 可逆. 进而 $\boldsymbol{A}+\delta\boldsymbol{A} = \boldsymbol{A}\left(\boldsymbol{E}+\boldsymbol{A}^{-1}\delta\boldsymbol{A}\right)$，可得 $\boldsymbol{A}+\delta\boldsymbol{A}$ 可逆.

(2) 由 $(\boldsymbol{A}+\delta\boldsymbol{A})^{-1} = \left(\boldsymbol{E}+\boldsymbol{A}^{-1}\delta\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{-1}$, 可得

$$\begin{aligned}\left\|\boldsymbol{A}^{-1}-(\boldsymbol{A}+\delta\boldsymbol{A})^{-1}\right\|_a &= \left\|\left[\boldsymbol{E}-\left(\boldsymbol{E}+\boldsymbol{A}^{-1}\delta\boldsymbol{A}\right)^{-1}\right]\boldsymbol{A}^{-1}\right\|_a \\ &= \left\|\boldsymbol{A}^{-1}\delta\boldsymbol{A}\left(\boldsymbol{E}+\boldsymbol{A}^{-1}\delta\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{-1}\right\|_a \\ &\leq \left\|\boldsymbol{A}^{-1}\delta\boldsymbol{A}\right\|_a\left\|\left(\boldsymbol{E}+\boldsymbol{A}^{-1}\delta\boldsymbol{A}\right)^{-1}\right\|_a\left\|\boldsymbol{A}^{-1}\right\|_a.\end{aligned}$$

结合定理 1，有

$$
\begin{aligned}
\frac{\left\|A^{-1}-(A+\delta A)^{-1}\right\|_a}{\left\|A^{-1}\right\|_a} &\leq \frac{\left\|A^{-1}\delta A\right\|_a}{1-\left\|A^{-1}\delta A\right\|_a} \\
&\leq \frac{\left\|A^{-1}\right\|_a\|\delta A\|_a}{1-\left\|A^{-1}\right\|_a\|\delta A\|_a} \\
&= \frac{\left\|A^{-1}\right\|_a\|A\|_a\cdot\frac{\|\delta A\|_a}{\|A\|_a}}{1-\left\|A^{-1}\right\|_a\|A\|_a\cdot\frac{\|\delta A\|_a}{\|A\|_a}} \\
&= \frac{K(A)\cdot\frac{\|\delta A\|_a}{\|A\|_a}}{1-K(A)\cdot\frac{\|\delta A\|_a}{\|A\|_a}},
\end{aligned}
$$

得证.

定理 3. 在方程组 $Ax=b$ 中，A 固定且可逆，令 $b\neq 0$ 且有小扰动，解方程

$$
A(x+\delta x)=b+\delta b,
$$

得

$$
\frac{\|\delta x\|_a}{\|x\|_a}\leq K(A)\frac{\|\delta b\|_a}{\|b\|_a}.
$$

证明： 由 $A(x+\delta x)=b+\delta b$，可得 $A(\delta x)=\delta b$，于是 $\delta x=A^{-1}\delta b$. 两边取范数有

$$
\|\delta x\|_a\leq\left\|A^{-1}\right\|_a\|\delta b\|_a.
$$

又 $Ax=b$，则 $\|b\|_a\leq\|A\|_a\|x\|_a$，移项有

$$
\frac{1}{\|x\|_a}\leq\frac{\|A\|_a}{\|b\|_a}.
$$

相乘既得结论.

例 2. 方程组

$$
\begin{pmatrix} 0.505 & 0.495 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 1 \\ 2 \end{pmatrix}
$$

的解为 $(1,1)^T$. 设 b 有扰动 $\delta b=(0,\varepsilon)^T$, 则方程组

$$
\begin{pmatrix} 0.505 & 0.495 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 1 \\ 2 \end{pmatrix}+\begin{pmatrix} 0 \\ \varepsilon \end{pmatrix}
$$

的解为 $(1-49.5\varepsilon,1+50.5\varepsilon)^T$. 计算易得

$$
\frac{\|\delta x\|_1}{\|x\|_1}=50\varepsilon\leq K(A)\frac{\|\delta b\|_1}{\|b\|_1}=301\frac{\varepsilon}{3}\approx 100\varepsilon.
$$

定理 4. 在 $Ax=b$ 中, b 固定且非零, 令可逆矩阵 A 有小扰动 δA, 则当 $\left\|A^{-1}\right\|_a\|\delta A\|_a<1$ 时，解方程组

$$(A+\delta A)x=b,$$

得

$$\frac{\|\delta x\|_a}{\|x\|_a}\leq\frac{K(A)\frac{\|\delta A\|_a}{\|A\|_a}}{1-K(A)\frac{\|\delta A\|_a}{\|A\|_a}}.$$

证明： 由定理 2 知，当 $\left\|A^{-1}\delta A\right\|_a<1$ 时，$A+\delta A$ 可逆，故方程组 $(A+\delta A)x=b$ 有唯一解，记 $x^*=x+\delta x$.

$$\begin{aligned}\delta x&=x^*-x\\&=(A+\delta A)^{-1}b-A^{-1}b\\&=\left(E+A^{-1}\delta A\right)^{-1}A^{-1}b-A^{-1}b\\&=\left[\left(E+A^{-1}\delta A\right)^{-1}-E\right]A^{-1}b\\&=-A^{-1}\delta A\left(E+A^{-1}\delta A\right)^{-1}A^{-1}b\\&=-A^{-1}\delta A\left(E+A^{-1}\delta A\right)^{-1}x,\end{aligned}$$

即

$$\|\delta x\|_a=\left\|-A^{-1}\delta A\left(E+A^{-1}\delta A\right)^{-1}x\right\|_a\leq\left\|A^{-1}\delta A\right\|_a\left\|\left(E+A^{-1}\delta A\right)^{-1}\right\|_a\|x\|_a.$$

再由定理 2, 得

$$\frac{\|\delta x\|_a}{\|x\|_a}\leq\frac{K(A)\frac{\|\delta A\|_a}{\|A\|_a}}{1-K(A)\frac{\|\delta A\|_a}{\|A\|_a}}.$$

例 3. 设

$$A=\begin{pmatrix}2&6\\2&6.00001\end{pmatrix}\quad b=\begin{pmatrix}8\\8.00001\end{pmatrix},$$

则方程组 $Ax=b$ 得解为 $(1,1)^T$. 若 A,b 均有小扰动

$$\delta A=\begin{pmatrix}0&0\\0&-0.00002\end{pmatrix}\quad \delta b=\begin{pmatrix}0\\0.00001\end{pmatrix},$$

则方程组 $(A+\delta A)x=b+\delta b$ 的解为 $(10,-2)^T$.

定理 5. 设 A 可逆，$0\neq b\in\mathbf{C}^n$，$\left\|A^{-1}\right\|_a\|\delta A\|_a<1$，方程组 $Ax=b$ 的解是 x，则方程组 $(A+\delta A)(x+\delta x)=b+\delta b$ 有唯一解 $x+\delta x$, 并且满足

$$\frac{\|\delta x\|_a}{\|x\|_a}\leq\frac{K(A)}{1-K(A)\frac{\|\delta A\|_a}{\|A\|_a}}\left(\frac{\|\delta A\|_a}{\|A\|_a}+\frac{\|\delta b\|_a}{\|b\|_a}\right).$$

作业：用 Matlab 实现以下计算,

例 4. 设

$$\boldsymbol{A}=\begin{pmatrix} \mathbf{2} & \mathbf{6} \\ \mathbf{2} & \mathbf{6.00001} \end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix} \mathbf{2} & \mathbf{6} \\ \mathbf{2} & \mathbf{8.00001} \end{pmatrix},\quad \boldsymbol{\Delta}=\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{-0.00002} \end{pmatrix},$$

求 $\boldsymbol{A}^{-1}$, $(\boldsymbol{A}+\boldsymbol{\Delta})^{-1}$, $(\boldsymbol{B}+\boldsymbol{\Delta})^{-1}$. *有什么发现?*

例 5. 设

$$\boldsymbol{A}=\begin{pmatrix} \mathbf{2} & \mathbf{6} \\ \mathbf{2} & \mathbf{6.00001} \end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix} \mathbf{2} & \mathbf{6} \\ \mathbf{2} & \mathbf{8.00001} \end{pmatrix},$$

求矩阵 $\boldsymbol{A}$, $\boldsymbol{B}$ 在算子 1 范数、2 范数和无穷范数意义下的条件数. 有什么发现?

§1 矩阵的三角分解

程光辉

2020 年 4 月 14 日

定义 1 设 $\boldsymbol{A}$ 为 $\boldsymbol{m\times n}$ 复 (实) 矩阵，如果 $\mathbf{rank}(\boldsymbol{A})=\boldsymbol{m}$，则称 $\boldsymbol{A}$ 为行满秩矩阵，记为 $\boldsymbol{A}\in\mathbf{C}_{\boldsymbol{m}}^{\boldsymbol{m\times n}}$ $(\mathbf{R}_{\boldsymbol{m}}^{\boldsymbol{m\times n}})$; 如果 $\mathbf{rank}(\boldsymbol{A})=\boldsymbol{n}$, 则称 $\boldsymbol{A}$ 为列满秩矩阵, 记为 $\boldsymbol{A}\in\mathbf{C}_{\boldsymbol{n}}^{\boldsymbol{m\times n}}$ $(\mathbf{R}_{\boldsymbol{n}}^{\boldsymbol{m\times n}})$.

定义 2 设 $\boldsymbol{U}$ 为 $\boldsymbol{m\times n}$ 复矩阵，如果 $\boldsymbol{UU^H}=\boldsymbol{E_m}$，则 $\boldsymbol{U}$ 为行向量单位正交矩阵，记为 $\boldsymbol{U}\in\mathbf{U}_{\boldsymbol{m}}^{\boldsymbol{m\times n}}$；如果 $\boldsymbol{U^HU}=\boldsymbol{E_n}$，则 $\boldsymbol{U}$ 为列向量组单位正交矩阵，记为 $\boldsymbol{U}\in\mathbf{U}_{\boldsymbol{n}}^{\boldsymbol{m\times n}}$.

定义 3 若矩阵

$$\boldsymbol{R}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix},$$

则称 $\boldsymbol{R}$ 为上三角矩阵;

若矩阵

$$\tilde{\boldsymbol{R}}=\begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 0 & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

则称 $\tilde{\boldsymbol{R}}$ 为单位上三角矩阵.

定义 4 若矩阵

$$\boldsymbol{L}=\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix},$$

则称 $\boldsymbol{L}$ 为下三角矩阵;

若矩阵

$$\tilde{L}=\begin{bmatrix}1 & 0 & \cdots & 0\\ a_{21} & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1} & a_{n2} & \cdots & 1\end{bmatrix},$$

则称 $\tilde{L}$ 为单位下三角矩阵.

根据三角矩阵的结构性质，有如下性质：

(1) 若上三角矩阵 R 可逆，则 R^{-1} 也是上三角矩阵，且对角元是 R 对角元的倒数.

(2) 两个上三角矩阵 R_1 和 R_2 的乘积 R_1R_2 也是上三角矩阵，且对角元是 R_1 和 R_2 对角元之积.

酉矩阵也有类似性质，如下：

(1) 酉矩阵 U 的逆 U^{-1} 也是酉矩阵.

(2) 酉矩阵 U_1 和 U_2 的乘积 U_1U_2 也是酉矩阵.

定理 1 设 $A\in C_n^{n\times n}$，则 A 可唯一地分解为

$$A=U_1R,$$

其中 U_1 是酉矩阵，R 是正线上三角复矩阵，或 A 可唯一分解为

$$A=LU_2,$$

其中 L 是正线下三角复矩阵，U_2 是酉矩阵.

证明： 把矩阵 A 按照列划分，即 $A=(a_1,\ a_2,\cdots,\ a_n)$. 因为 $A\in C_n^{n\times n}$ 列满秩，则 a_1，$a_2,\cdots$，a_n 线性无关，对其进行 Gram-Schmidt 正交化和单位化，有

$$\begin{cases}b_1=\frac{a_1}{\|a_1\|}\\ b_i=\frac{a_i-\sum\limits_{j=1}^{i-1}(a_i,b_j)b_j}{\|a_i-\sum\limits_{j=1}^{i-1}(a_i,b_j)b_j\|}, \quad i=2,\cdots,n.\end{cases}$$

记 $r_{11}=\|a_1\|$，$r_{ii}=\|a_i-\sum\limits_{j=1}^{i-1}(a_i,b_j)b_j\|$ 和 $r_{ji}=(a_i,b_j)$，则有 $a_i=\sum\limits_{j=1}^{i}r_{ji}b_j$， $i=1,2,\cdots,n$.

于是有

$$A=(r_{11}b_1,r_{12}b_1+r_{22}b_2,\cdots,\sum_{j=1}^{n}r_{jn}b_j)$$

$$
\begin{aligned}
&= (b_1, b_2, \cdots, b_n)R \\
&= U_1 R,
\end{aligned}
$$

其中

$$
R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix}.
$$

唯一性：设 $A = U_1R_1 = U_2R_2$，其中 U_1, U_2 是酉矩阵，R_1, R_2 是上三角矩阵. 于是有

$$
U_2^H U_1 = R_2 R_1^{-1},
$$

既是酉矩阵又是上三角矩阵，即为对角矩阵。又根据对角元，知为单位矩阵，即 $R_1 = R_2$，$U_1 = U_2$，唯一性得证.

推论 1 *(矩阵的 QR 分解)* 设 $A \in \mathbb{R}_n^{n\times n}$，则 A 可唯一地分解为

$$
A = Q_1 R,
$$

其中 Q_1 是正交矩阵，R 是正线上三角实矩阵；或 A 可唯一分解为

$$
A = LQ_2,
$$

其中 L 是正线下三角实矩阵，Q_2 是正交矩阵.

例 1 求三阶实矩阵 $A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 1 & 2 & 3 \end{bmatrix}$ 的 QR 分解.

解：令 $A = (a_1, a_2, a_3)$，对 a_1, a_2, a_3 使用 Gram-Schmidt 正交化得

$$
\begin{aligned}
b_1 &= a_1 = (1, 1, 1)^T \\
b_2 &= a_2 - \frac{(a_2, b_1)}{(b_1, b_1)} b_1 = a_2 - b_1 = (-1, 0, 1)^T \\
b_3 &= a_3 - \frac{(a_3, b_2)}{(b_2, b_2)} b_2 - \frac{(a_3, b_1)}{(b_1, b_1)} b_1 = a_3 - \frac{1}{2} b_2 - \frac{5}{3} b_1 = \frac{5}{6}(1, -2, 1)^T.
\end{aligned}
$$

单位化有

$$
\begin{aligned}
\gamma_1 &= \frac{b_1}{\|b_1\|} = \frac{1}{\sqrt{3}}(1, 1, 1)^T \\
\gamma_2 &= \frac{b_2}{\|b_2\|} = \frac{1}{\sqrt{2}}(-1, 0, 1)^T \\
\gamma_3 &= \frac{b_3}{\|b_3\|} = \frac{1}{\sqrt{6}}(1, -2, 1)^T.
\end{aligned}
$$

综上，有

$$
\begin{aligned}
\boldsymbol{A} &= (\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3) \begin{bmatrix} 1 & 1 & \frac{5}{3} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} \\
&= (\gamma_1, \gamma_2, \gamma_3) \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \frac{5}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 1 & 1 & \frac{5}{3} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} \\
&= \boldsymbol{Q} \begin{bmatrix} \sqrt{3} & \sqrt{3} & \frac{5}{\sqrt{3}} \\ 0 & \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{5}{\sqrt{6}} \end{bmatrix}.
\end{aligned}
$$

推论 2 设 $\boldsymbol{A}$ 是实对称正定矩阵，则存在唯一的正线上三角实矩阵 $\boldsymbol{R}$，使

$$\boldsymbol{A} = \boldsymbol{R}^T \boldsymbol{R}.$$

证明：因为 $\boldsymbol{A}$ 是实对称正定矩阵，则存在实可逆矩阵 $\boldsymbol{P}$ 使得

$$\boldsymbol{A} = \boldsymbol{P}^T \boldsymbol{P}.$$

又因推论 1 知，$\boldsymbol{P} = \boldsymbol{QR}$，其中 $\boldsymbol{Q}$ 是实正交矩阵，$\boldsymbol{R}$ 是正线上三角实矩阵. 进而有

$$\boldsymbol{A} = \boldsymbol{R}^T \boldsymbol{Q}^T \boldsymbol{Q} \boldsymbol{R} = \boldsymbol{R}^T \boldsymbol{R}.$$

(唯一性) 设 $\boldsymbol{A} = \boldsymbol{R}_1^T \boldsymbol{R}_1 = \boldsymbol{R}_2^T \boldsymbol{R}_2$，其中 $\boldsymbol{R}_1$，$\boldsymbol{R}_2$ 正线上三角实矩阵，则

$$(\boldsymbol{R}_1^T)^{-1} \boldsymbol{R}_2^T = \boldsymbol{R}_1 (\boldsymbol{R}_2)^{-1},$$

上式等号左侧为正线下三角矩阵，等号右侧为上三角矩阵，即为对角矩阵. 再根据对角元素，知为单位矩阵，即有

$$\boldsymbol{R}_1 = \boldsymbol{R}_2,$$

得证.

推论 3 设 $\boldsymbol{A}$ 是正定 **Hermitian** 矩阵，则存在唯一的正线上三角复矩阵 $\boldsymbol{R}$，使

$$\boldsymbol{A} = \boldsymbol{R}^H \boldsymbol{R}.$$

定理 2 设 $\boldsymbol{A} \in \mathbf{C}_n^{n \times n}$，用 $\boldsymbol{L}$，$\tilde{\boldsymbol{L}}$，$\boldsymbol{R}$，$\tilde{\boldsymbol{R}}$ 和 $\boldsymbol{D}$ 分别表示下三角复矩阵、单位下三角复矩阵、上三角复矩阵、单位上三角复矩阵和对角矩阵，则下列命题等价：

(1) $\boldsymbol{\Delta}_k = \det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix} \neq 0,$

(2) $\boldsymbol{A}$ 可唯一地分解为 $\boldsymbol{A}=\boldsymbol{L}\tilde{\boldsymbol{R}}$,

(3) $\boldsymbol{A}$ 可唯一地分解为 $\boldsymbol{A}=\tilde{\boldsymbol{L}}\boldsymbol{R}$,

(4) $\boldsymbol{A}$ 可唯一地分解为 $\boldsymbol{A}=\tilde{\boldsymbol{L}}\boldsymbol{D}\tilde{\boldsymbol{R}}$.

证明： $(1)\Rightarrow(2)$ 利用数学归纳法进行证明.

(i) 若 $\boldsymbol{A}$ 为一阶方阵，则显然有 $\boldsymbol{A}=(a_{11})(1)=\boldsymbol{L}\tilde{\boldsymbol{R}}$.

(ii) 若 $\boldsymbol{A}$ 为 $n-1$ 阶方阵，假设有 $\boldsymbol{A}=\boldsymbol{L}_{n-1}\tilde{\boldsymbol{R}}_{n-1}$ 成立.

(iii) 若 $\boldsymbol{A}$ 为 n 阶方阵，则有

$$
\begin{aligned}
\boldsymbol{A}&=\begin{bmatrix}\boldsymbol{A}_{n-1} & \boldsymbol{\beta}\\ \boldsymbol{\alpha}^T & a_{nn}\end{bmatrix}\\
&=\begin{bmatrix}\boldsymbol{A}_{n-1} & 0\\ \boldsymbol{\alpha}^T & a_{nn}-\boldsymbol{\alpha}^T\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta}\end{bmatrix}\begin{bmatrix}\boldsymbol{E}_{n-1} & \boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta}\\ 0 & 1\end{bmatrix}\\
&=\begin{bmatrix}\boldsymbol{L}_{n-1}\tilde{\boldsymbol{R}}_{n-1} & 0\\ \boldsymbol{\alpha}^T & a_{nn}-\boldsymbol{\alpha}^T\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta}\end{bmatrix}\begin{bmatrix}\boldsymbol{E}_{n-1} & \boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta}\\ 0 & 1\end{bmatrix}\\
&=\begin{bmatrix}\boldsymbol{L}_{n-1} & 0\\ \boldsymbol{\alpha}^T\tilde{\boldsymbol{R}}_{n-1}^{-1} & a_{nn}-\boldsymbol{\alpha}^T\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta}\end{bmatrix}\begin{bmatrix}\tilde{\boldsymbol{R}}_{n-1} & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\boldsymbol{E}_{n-1} & \boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta}\\ 0 & 1\end{bmatrix}\\
&=\begin{bmatrix}\boldsymbol{L}_{n-1} & 0\\ \boldsymbol{\alpha}^T\tilde{\boldsymbol{R}}_{n-1}^{-1} & a_{nn}-\boldsymbol{\alpha}^T\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta}\end{bmatrix}\begin{bmatrix}\tilde{\boldsymbol{R}}_{n-1} & \tilde{\boldsymbol{R}}_{n-1}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta}\\ 0 & 1\end{bmatrix}\\
&=\boldsymbol{L}\tilde{\boldsymbol{R}},
\end{aligned}
$$

得证.

$(2)\Rightarrow(1)$ 由 $\boldsymbol{A}=\boldsymbol{L}\tilde{\boldsymbol{R}}$，分块得

$$
\begin{aligned}
\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix}&=\begin{bmatrix}\boldsymbol{L}_{11} & 0\\ \boldsymbol{L}_{21} & \boldsymbol{L}_{22}\end{bmatrix}\begin{bmatrix}\tilde{\boldsymbol{R}}_{11} & \tilde{\boldsymbol{R}}_{12}\\ 0 & \tilde{\boldsymbol{R}}_{22}\end{bmatrix}\\
&=\begin{bmatrix}\boldsymbol{L}_{11}\tilde{\boldsymbol{R}}_{11} & \boldsymbol{L}_{11}\tilde{\boldsymbol{R}}_{12}\\ \boldsymbol{L}_{21}\tilde{\boldsymbol{R}}_{11} & \boldsymbol{L}_{21}\tilde{\boldsymbol{R}}_{12}+\boldsymbol{L}_{22}\tilde{\boldsymbol{R}}_{22}\end{bmatrix}.
\end{aligned}
$$

有 $\boldsymbol{A}_{11}=\boldsymbol{L}_{11}\tilde{\boldsymbol{R}}_{11}$，进而

$$
\Delta_k=\det(\boldsymbol{A}_{11})=\det(\boldsymbol{L}_{11})\det(\tilde{\boldsymbol{R}}_{11})=\det(\boldsymbol{L}_{11})=l_{11}l_{22}\cdots l_{kk}\neq 0.
$$

例 2 求 $\boldsymbol{A}=\begin{bmatrix}2 & -1 & 3\\ 1 & 2 & 1\\ 2 & 4 & 3\end{bmatrix}$ 的 $\tilde{\boldsymbol{L}}\boldsymbol{R}$ 及 $\tilde{\boldsymbol{L}}\boldsymbol{D}\tilde{\boldsymbol{R}}$ 分解.

解：计算得 $\boldsymbol{\Delta_1=2}$，$\boldsymbol{\Delta_2=5}$，$\boldsymbol{\Delta_3=5}$，则存在所求分解.

令

$$\tilde{L}=\begin{bmatrix}1&0&0\\l_{21}&1&0\\l_{31}&l_{32}&1\end{bmatrix},\quad R=\begin{bmatrix}r_{11}&r_{12}&r_{13}\\0&r_{22}&r_{23}\\0&0&r_{33}\end{bmatrix},$$

则

$$\tilde{L}R=\begin{bmatrix}r_{11}&r_{12}&r_{13}\\l_{21}r_{11}&l_{21}r_{12}+r_{22}&l_{21}r_{13}+r_{23}\\l_{31}r_{11}&l_{31}r_{12}+l_{32}r_{22}&l_{31}r_{13}+l_{32}r_{23}+l_{33}r_{33}\end{bmatrix}=\begin{bmatrix}2&-1&3\\1&2&1\\2&4&3\end{bmatrix}=A,$$

利用追赶法求解得

$$\tilde{L}=\begin{bmatrix}1&0&0\\0.5&1&0\\1&2&1\end{bmatrix},\ R=\begin{bmatrix}2&-1&3\\0&2.5&-0.5\\0&0&1\end{bmatrix},\ D=\begin{bmatrix}2&0&0\\0&2.5&0\\0&0&1\end{bmatrix},\ \tilde{R}=\begin{bmatrix}1&-0.5&1.5\\0&1&-0.2\\0&0&1\end{bmatrix}.$$

定理 3 设 $A\in\mathbf{C}_n^{m\times n}$，则存在 m 阶酉矩阵 U 和 n 阶正线上三角复矩阵 R，使得

$$A=U\begin{bmatrix}R\\0\end{bmatrix}.$$

设 $A\in\mathbf{C}_m^{m\times n}$，则存在 n 阶酉矩阵 U 和 m 阶正线下三角复矩阵 L，使得

$$A=\begin{bmatrix}L&0\end{bmatrix}U.$$

证明：因为 $A\in\mathbf{C}_n^{m\times n}$，则列向量组 $a_1,a_2,\cdots,a_n$ 线性无关，添加 $m-n$ 个向量使得

$$a_1,a_2,\cdots,a_n,a_{n+1},\cdots,a_m$$

线性无关 ($\mathbf{C}^m$ 空间的一组基).

类似于定理 1 的证明，对其进行 **Gram-Schmidt** 正交化得

$$\begin{aligned}A&=(a_1,a_2,\cdots,a_n)\\&=(r_{11}b_1,r_{12}b_1+r_{22}b_2,\cdots,\sum_{j=1}^{n}r_{jn}b_j)\\&=(b_1,b_2,\cdots,b_n,\cdots,b_m)\begin{bmatrix}r_{11}&r_{12}&\cdots&r_{1n}\\0&r_{22}&\cdots&r_{2n}\\\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&r_{nn}\\0&0&\cdots&0\\\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&0\end{bmatrix}\end{aligned}$$

$$
= \boldsymbol{U}\begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{0} \end{bmatrix},
$$

得证.

定理 4 设 $\boldsymbol{A} \in \mathbf{C}_n^{m\times n}$，则 $\boldsymbol{A}$ 可唯一地分解为

$$
\boldsymbol{A} = \boldsymbol{U}\boldsymbol{R},
$$

其中 $\boldsymbol{U} \in \mathbf{U}_n^{m\times n}$，$\boldsymbol{R}$ 是 n 阶正线上三角矩阵；设 $\boldsymbol{A} \in \mathbf{C}_m^{m\times n}$，则 $\boldsymbol{A}$ 可唯一地分解为

$$
\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U},
$$

其中 $\boldsymbol{L}$ 是 m 阶正线下三角矩阵，$\boldsymbol{U} \in \mathbf{U}_m^{m\times n}$.

证明： (可利用定理 3 的结论证明存在性.)

因为 $\boldsymbol{A} \in \mathbf{C}_n^{m\times n}$，对 $\forall \boldsymbol{x} \neq \boldsymbol{0}$，有 $\boldsymbol{A}\boldsymbol{x} \neq \boldsymbol{0}$，故

$$
\boldsymbol{x}^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{A}\boldsymbol{x})^H(\boldsymbol{A}\boldsymbol{x}) > 0,
$$

则 $\boldsymbol{A}^H\boldsymbol{A}$ 为正定 **Hermitian** 矩阵. 进而存在正线上三角复矩阵 $\boldsymbol{R}$，使得 $\boldsymbol{A}^H\boldsymbol{A} = \boldsymbol{R}^H\boldsymbol{R}$.

令 $\boldsymbol{U} = \boldsymbol{A}\boldsymbol{R}^{-1}$，则 $\boldsymbol{U}^H\boldsymbol{U} = (\boldsymbol{A}\boldsymbol{R}^{-1})^H(\boldsymbol{A}\boldsymbol{R}^{-1}) = (\boldsymbol{R}^H)^{-1}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{R}^{-1} = \boldsymbol{E}_n$，即 $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{R}$ 得证.

(唯一性) 令 $\boldsymbol{A} = \boldsymbol{U}_1\boldsymbol{R}_1 = \boldsymbol{U}_2\boldsymbol{R}_2$，则

$$
\boldsymbol{A}^H\boldsymbol{A} = \boldsymbol{R}_1^H\boldsymbol{U}_1^H\boldsymbol{U}_1\boldsymbol{R}_1 = \boldsymbol{R}_1^H\boldsymbol{R}_1 = \boldsymbol{R}_2^H\boldsymbol{R}_2,
$$

根据定理 1 的推论 3 知，$\boldsymbol{R}_1 = \boldsymbol{R}_2$，进而 $\boldsymbol{U}_1 = \boldsymbol{U}_2$，得证.

另证： $\boldsymbol{U}_2^H\boldsymbol{U}_1 = \boldsymbol{R}_2\boldsymbol{R}_1^{-1}$，$\boldsymbol{U}_1^H\boldsymbol{U}_2 = \boldsymbol{R}_1\boldsymbol{R}_2^{-1}$，易知上述两个等式左侧矩阵互为共轭转置，有

$$
\boldsymbol{U}_2^H\boldsymbol{U}_1 = \boldsymbol{R}_2\boldsymbol{R}_1^{-1} = (\boldsymbol{R}_1\boldsymbol{R}_2^{-1})^H,
$$

上式第二个等号左右两侧分别为上三角矩阵和下三角矩阵，即为对角矩阵. 再根据对角元易证 $\boldsymbol{R}_1 = \boldsymbol{R}_2$，进而 $\boldsymbol{U}_1 = \boldsymbol{U}_2$，得证.

定理 5 设 $\boldsymbol{A} \in \mathbf{C}_r^{m\times n}$，则存在酉矩阵 $\boldsymbol{U} \in \mathbf{U}^{m\times m}$ 和 $\boldsymbol{V} \in \mathbf{U}^{n\times n}$ 及 r 阶正线下三角矩阵 $\boldsymbol{L}$，使得

$$
\boldsymbol{A} = \boldsymbol{U}\begin{bmatrix} \boldsymbol{L} & 0 \\ 0 & 0 \end{bmatrix}\boldsymbol{V}.
$$

证明： 因为 $\boldsymbol{A} \in \mathbf{C}_r^{m\times n}$，则存在酉矩阵 $\boldsymbol{P}$ (列初等矩阵的乘积)，使得

$$
\boldsymbol{A}\boldsymbol{P} = (\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n),
$$

其中 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_r$ 线性无关. 于是存在矩阵 $\boldsymbol{C}$, 使得

$$(\boldsymbol{\alpha}_{r+1}, \cdots, \boldsymbol{\alpha}_n) = (\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_r)\boldsymbol{C}.$$

再根据定理 3，有

$$\begin{aligned} \boldsymbol{AP} &= (\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_r)\begin{bmatrix} \boldsymbol{E}_r & \boldsymbol{C} \end{bmatrix} \\ &= \boldsymbol{U}\begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{E}_r & \boldsymbol{C} \end{bmatrix} \\ &= \boldsymbol{U}\begin{bmatrix} \boldsymbol{R} & \boldsymbol{RC} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}. \end{aligned}$$

令 $\boldsymbol{B} = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{RC} \end{bmatrix} \in \mathbb{C}_r^{r\times n}$，则 $\boldsymbol{B} = \begin{bmatrix} \boldsymbol{L} & \boldsymbol{0} \end{bmatrix}\boldsymbol{V}_1$.

于是有

$$\boldsymbol{A} = \boldsymbol{U}\begin{bmatrix} \boldsymbol{R} & \boldsymbol{RC} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}\boldsymbol{P}^{-1} = \boldsymbol{U}\begin{bmatrix} \boldsymbol{L} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}\boldsymbol{V}_1\boldsymbol{P}^{-1} = \boldsymbol{U}\begin{bmatrix} \boldsymbol{L} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}\boldsymbol{V},$$

其中 $\boldsymbol{V} = \boldsymbol{V}_1\boldsymbol{P}^{-1}$.

§2 矩阵的谱分解

程光辉

2020 年 4 月 4 日

定义 1 设 $\boldsymbol{\lambda_1, \lambda_2, \cdots, \lambda_k}$ 是 $\boldsymbol{A} \in \mathbf{C}^{n\times n}$ 的 $\boldsymbol{k}$ 个互异特征值,其重数分别为 $\boldsymbol{r_1, r_2, \cdots, r_k}$,则称 $\boldsymbol{r_i}$ 为矩阵 $\boldsymbol{A}$ 的特征值 $\boldsymbol{\lambda_i}$ 的代数重复度. 齐次方程组

$$\boldsymbol{Ax = \lambda_i x \quad (i = 1, 2, \cdots, k)}$$

的解空间 $\boldsymbol{V_{\lambda_i}}$ 称为 $\boldsymbol{A}$ 对应于特征值 $\boldsymbol{\lambda_i}$ 的特征子空间，$\boldsymbol{V_{\lambda_i}}$ 的维数称为 $\boldsymbol{A}$ 的特征值的几何重复度.

定理 1 矩阵 $\boldsymbol{A}$ 的任意特征值的几何重复度不大于它的代数重复度.

定义 2 若矩阵 $\boldsymbol{A}$ 的每个特征值的几何重复度与它的代数重复度相等，则称 $\boldsymbol{A}$ 为单纯矩阵.

定理 2 矩阵 $\boldsymbol{A}$ 是单纯矩阵的充要条件是 $\boldsymbol{A}$ 与对角矩阵相似.

定理 3 设矩阵 $\boldsymbol{A} \in \mathbf{C}^{n\times n}$ 是单纯矩阵,则 $\boldsymbol{A}$ 可分解为一系列幂等矩阵 $\boldsymbol{A_i}$ $\boldsymbol{(i = 1, 2, \cdots, n)}$ 的加权和，即

$$\boldsymbol{A = \sum_{i=1}^{n} \lambda_i A_i},$$

其中 $\boldsymbol{\lambda_i}$ $\boldsymbol{(i = 1, 2, \cdots, n)}$ 是 $\boldsymbol{A}$ 的特征值.

证明：因为 $\boldsymbol{A}$ 是单纯矩阵，则存在可逆矩阵 $\boldsymbol{P}$ 使得

$$\boldsymbol{A = P\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)P^{-1}},$$

其中 $\boldsymbol{\lambda_i}$ 是 $\boldsymbol{A}$ 的特征值.

令 $\boldsymbol{P = (v_1, v_2, \cdots, v_n)}$，$\boldsymbol{P^{-1} = (w_1, w_2, \cdots, w_n)^T}$，则有

$$\boldsymbol{A = (v_1, v_2, \cdots, v_n)\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}\begin{bmatrix} w_1^T \\ w_2^T \\ \vdots \\ w_n^T \end{bmatrix} = \sum_{i=1}^{n} \lambda_i A_i},$$

其中 $\boldsymbol{A}_i = \boldsymbol{v}_i\boldsymbol{w}_i^T$.

又因为 $\boldsymbol{P}^{-1}\boldsymbol{P} = \boldsymbol{E}_n$，可知

$$\boldsymbol{w}_i^T\boldsymbol{v}_j = \begin{cases} 1, & j = i, \\ 0, & j \neq i, \end{cases}$$

进而有

$$\boldsymbol{A}_i\boldsymbol{A}_j = \boldsymbol{v}_i\boldsymbol{w}_i^T\boldsymbol{v}_j\boldsymbol{w}_j^T = \begin{cases} \boldsymbol{v}_i\boldsymbol{w}_i^T = \boldsymbol{A}_i, & j = i, \\ \boldsymbol{O}, & j \neq i, \end{cases}$$

故 $\boldsymbol{A}_i$ 为幂等矩阵，得证.

$\boldsymbol{A}_i$ 的性质：

(1) 幂等性：$\boldsymbol{A}_i^2 = \boldsymbol{A}_i$

(2) 分离性：$\boldsymbol{A}_i\boldsymbol{A}_j = \boldsymbol{O} \quad (i \neq j)$

(3) 可加性：$\sum\limits_{i=1}^{n} \boldsymbol{A}_i = \boldsymbol{E}_n$

证明： (3)

$$\boldsymbol{P}\boldsymbol{P}^{-1} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n)\begin{bmatrix} \boldsymbol{w}_1^T \\ \boldsymbol{w}_2^T \\ \vdots \\ \boldsymbol{w}_n^T \end{bmatrix} = \sum_{i=1}^{n} \boldsymbol{v}_i\boldsymbol{w}_i^T = \sum_{i=1}^{n} \boldsymbol{A}_i = \boldsymbol{E}_n.$$

定理 4 设 $\boldsymbol{A} \in \mathbf{C}^{n\times n}$，它有 k 个相异特征值 λ_i $(i = 1, 2, \cdots, k)$，则 $\boldsymbol{A}$ 是单纯矩阵的充要条件是存在 k 个矩阵 $\boldsymbol{A}_i$ $(i = 1, 2, \cdots, k)$ 满足

(1) $\boldsymbol{A}_i\boldsymbol{A}_j = \begin{cases} \boldsymbol{A}_i, & j = i, \\ \boldsymbol{O}, & j \neq i, \end{cases}$,

(2) $\sum\limits_{i=1}^{k} \boldsymbol{A}_i = \boldsymbol{E}_n$,

(3) $\boldsymbol{A} = \sum\limits_{i=1}^{k} \lambda_i\boldsymbol{A}_i$.

定义 3 若 n 阶复矩阵 $\boldsymbol{A}$ 满足

$$\boldsymbol{A}\boldsymbol{A}^H = \boldsymbol{A}^H\boldsymbol{A},$$

则称矩阵 $\boldsymbol{A}$ 为正规矩阵.

思考：哪些矩阵属于正规矩阵？

引理 1 若 $\boldsymbol{A}$ 为正规矩阵，且 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 酉相似，则 $\boldsymbol{B}$ 为正规矩阵.

证明： 因为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 酉相似，则存在酉矩阵 $\boldsymbol{U}$，使得 $\boldsymbol{B}=\boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{U}$，进而

$$\begin{aligned}
\boldsymbol{B}\boldsymbol{B}^{H} &= (\boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{U})(\boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{U})^{H}\\
&= \boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{U}\boldsymbol{U}^{H}\boldsymbol{A}^{H}\boldsymbol{U}\\
&= \boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{A}^{H}\boldsymbol{U}\\
&= \boldsymbol{U}^{H}\boldsymbol{A}^{H}\boldsymbol{A}\boldsymbol{U} \quad (\ \boldsymbol{A}\text{为正规矩阵})\\
&= \boldsymbol{U}^{H}\boldsymbol{A}^{H}\boldsymbol{U}\boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{U}\\
&= (\boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{U})^{H}(\boldsymbol{U}^{H}\boldsymbol{A}\boldsymbol{U})\\
&= \boldsymbol{B}^{H}\boldsymbol{B},
\end{aligned}$$

得证.

引理 2 *(Schur 分解)* 若 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，则存在酉矩阵 $\boldsymbol{U}$，使得

$$\boldsymbol{A}=\boldsymbol{U}\boldsymbol{R}\boldsymbol{U}^{H},$$

其中 $\boldsymbol{R}$ 是一个上三角矩阵，且对角元为矩阵 $\boldsymbol{A}$ 的特征值.

证明： 对任意的 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 都与 Jordan 标准形相似，即存在可逆矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{A}=\boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1}.$$

因为 $\boldsymbol{P}$ 可逆，存在酉矩阵 $\boldsymbol{U}$ 和正线上三角矩阵 $\boldsymbol{R}_1$，使得 $\boldsymbol{P}=\boldsymbol{U}\boldsymbol{R}_1$.

　　于是有

$$\boldsymbol{A}=\boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1}=\boldsymbol{U}\boldsymbol{R}_1\boldsymbol{J}\boldsymbol{R}_1^{-1}\boldsymbol{U}^{-1}=\boldsymbol{U}\boldsymbol{R}\boldsymbol{U}^{H},$$

其中 $\boldsymbol{R}=\boldsymbol{R}_1\boldsymbol{J}\boldsymbol{R}_1^{-1}$ 为上三角矩阵，故得证.

引理 3 设 $\boldsymbol{A}$ 是三角矩阵，则 $\boldsymbol{A}$ 是正规矩阵充要条件是 $\boldsymbol{A}$ 对角矩阵.

证明： 不妨设

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ 0 & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & a_{nn}\end{bmatrix}.$$

(必要性) 因为 $\boldsymbol{A}$ 是正规矩阵，则

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ 0 & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & a_{nn}\end{bmatrix}\begin{bmatrix} \bar{a}_{11} & 0 & \cdots & 0\\ \bar{a}_{12} & \bar{a}_{22} & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots\\ \bar{a}_{1n} & \bar{a}_{2n} & \cdots & \bar{a}_{nn}\end{bmatrix}=\begin{bmatrix} \bar{a}_{11} & 0 & \cdots & 0\\ \bar{a}_{12} & \bar{a}_{22} & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots\\ \bar{a}_{1n} & \bar{a}_{2n} & \cdots & \bar{a}_{nn}\end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ 0 & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & a_{nn}\end{bmatrix},$$

比较上式两端矩阵第一行第一列的元素，有

$$\sum_{i=1}^{n} a_{1i}\bar{a}_{1i} = \sum_{i=1}^{n} |a_{1i}|^2 = a_{11}\bar{a}_{11} = |a_{11}|^2,$$

即有

$$a_{1i} = 0, \quad i = 2, 3, \cdots, n.$$

依此类推，即可得证.

(充分性) 直接验证，显然成立.

定理 5 n 阶复矩阵 A 是正规矩阵的充要条件是 A 与对角矩阵酉相似，即存在 n 阶酉矩阵 U，使得

$$A = U\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)U^H,$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 A 的 n 个特征值.

证明：(必要性) 由引理 2，有 $A = URU^H$. 又因为 A 是正规矩阵及引理 1 和引理 3，有 R 为对角矩阵.

(充分性) 直接由引理 1 得证.

定理 6 设 $A \in \mathrm{C}^{n\times n}$，它有 k 个相异特征值 λ_i $(i = 1, 2, \cdots, k)$，则 A 是正规矩阵的充要条件是存在 k 个矩阵 A_i $(i = 1, 2, \cdots, k)$ 满足

(1) $A_iA_j = \begin{cases} A_i, & j = i, \\ O, & j \neq i, \end{cases}$,

(2) $\sum\limits_{i=1}^{k} A_i = E_n$,

(3) $A = \sum\limits_{i=1}^{k} \lambda_i A_i$,

(4) $A_i^H = A_i$.

证明：(必要性) 因为 A 是正规矩阵，则

$$\begin{aligned}
A &= U\mathrm{diag}(\lambda_1E_{r_1}, \lambda_2E_{r_2}, \cdots, \lambda_kE_{r_k})U^H \\
&= (U_1, U_2, \cdots, U_k)\begin{bmatrix} \lambda_1E_{r_1} & 0 & \cdots & 0 \\ 0 & \lambda_2E_{r_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_kE_{r_k} \end{bmatrix}\begin{bmatrix} U_1^H \\ U_2^H \\ \vdots \\ U_k^H \end{bmatrix} \\
&= \sum_{i=1}^{k} \lambda_i U_i U_i^H \\
&= \sum_{i=1}^{k} \lambda_i A_i,
\end{aligned}$$

其中 $A_i = U_iU_i^H$.

根据 $UU^H = U^HU = E_n$，有

$$U_iU_j^H = \begin{cases} E_{r_i}, & j = i, \\ O, & j \neq i. \end{cases}$$

于是有

$$A_iA_j = U_iU_i^HU_jU_j^H = \begin{cases} U_iU_i^H = A_i, & j = i, \\ O, & j \neq i, \end{cases}$$

$$\sum_{i=1}^{k} A_i = \sum_{i=1}^{k} U_iU_i^H = UU^H = E_n,$$

$$A_i^H = (U_iU_i^H)^H = A_i.$$

(充分性) 只需验证 $AA^H = A^HA$ 即可.

$$\begin{aligned} AA^H &= \left(\sum_{i=1}^{k} \lambda_i A_i\right)\left(\sum_{j=1}^{k} \lambda_j A_j\right)^H \\ &= \sum_{i=1}^{k}\sum_{j=1}^{k} \lambda_i \bar{\lambda}_j A_i A_j^H \\ &= \sum_{i=1}^{k}\sum_{j=1}^{k} \lambda_i \bar{\lambda}_j A_i A_j \\ &= \sum_{i=1}^{k} |\lambda_i|^2 A_i. \end{aligned}$$

同理，可得

$$A^HA = \sum_{i=1}^{k} |\lambda_i|^2 A_i,$$

故 A 是正规矩阵.

§3 **Hermitian** 矩阵及其分解

程光辉

2020 年 4 月 4 日

定义 1 设 $A \in \mathbf{C}^{n\times n}$，若 $A^H = A$，则称 A 为 **Hermitian** 矩阵；若 $A^H = -A$，则称 A 为反 **Hermitian** 矩阵.

定理 1 设 $A \in \mathbf{C}^{n\times n}$ 是 **Hermitian** 矩阵，则

(1) $(A\alpha, \beta) = (\alpha, A\beta), \forall \alpha, \beta \in \mathbf{C}^n$*,*

(2) A *的特征值都是实数,*

(3) 属于 A *不同特征值的特征向量正交,*

(4) A *与矩阵* $\begin{bmatrix} E_p & 0 & 0 \\ 0 & -E_{r-p} & 0 \\ 0 & 0 & 0 \end{bmatrix}$ *合同，其中* $\mathrm{rank}(A) = r$*，*p *为正惯性指数,*

(5) $U^H A U = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$*，其中* U *为酉矩阵.*

定义 2 设 $A \in \mathbf{C}^{n\times n}$ 是 **Hermitian** 矩阵，对任意非零向量 $x \in \mathbf{C}^n$，都有

$$f(x) = x^H A x > 0 \quad (\geq 0),$$

则称二次型 $f(x)$ 是正定 *(*半正定*)* 二次型，此时系数矩阵 A 称为正定 *(*半正定*)* 矩阵.

定理 2 设 $A = (a_{ij}) \in \mathbf{C}^{n\times n}$ 是正定 **Hermitian** 矩阵，则

(1) $a_{ii} > 0, i = 1, 2, \cdots, n$*;*

(2) A *的特征值* $\lambda_i > 0, i = 1, 2, \cdots, n$*;*

(3) 存在正定矩阵 B*，使得* $A = B^k$*，*$k \in Z^+$*;*

(4) 存在正线下三角矩阵 L*，使得* $A = LL^H$*, (***Cholesky** *分解);*

(5) $\det(A) \leq a_{11}a_{22}\cdots a_{nn} = \Pi_{i=1}^n a_{ii}$*, (***Hadamard** *不等式);*

(6) $\boldsymbol{A}$ *与单位矩阵合同;*

(7) $\boldsymbol{A}$ *的顺序主子式都为正.*

证明： (5) 由 (4) 知，存在正线下三角 $\boldsymbol{L} = (l_{ij})$，使得 $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^H$，则矩阵 $\boldsymbol{A}$ 的对角元素为

$$a_{ii} = \sum_{k=1}^{i} l_{ik}\bar{l}_{ik} = \sum_{k=1}^{i} |l_{ik}|^2 \geq |l_{ii}|^2, \quad i = 1, 2, \cdots n.$$

又因为

$$\det(A) = \det(LL^H) = \det(L)\det(L^H) = \Pi_{i=1}^n l_{ii}^2 \leq \Pi_{i=1}^n a_{ii},$$

得证.

定理 3 *设* $\boldsymbol{A} \in \mathbf{C}^{n\times n}$ *是正定* **Hermitian** *矩阵,* $\boldsymbol{B} \in \mathbf{C}^{n\times n}$ *是* **Hermitian** *矩阵, 则存在可逆矩阵* $\boldsymbol{T} \in \mathbf{C}^{n\times n}$*, 使得*

$$\boldsymbol{T}^H\boldsymbol{A}\boldsymbol{T} = \boldsymbol{E}_n, \quad \boldsymbol{T}^H\boldsymbol{B}\boldsymbol{T} = \boldsymbol{\Lambda},$$

其中 $\boldsymbol{\Lambda}$ *为对角矩阵.*

证明： 因为 $\boldsymbol{A} \in \mathbf{C}^{n\times n}$ 是正定 **Hermitian** 矩阵，即存在可逆矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{P}^H\boldsymbol{A}\boldsymbol{P} = \boldsymbol{E}_n.$$

令 $\boldsymbol{Q} = \boldsymbol{P}^H\boldsymbol{B}\boldsymbol{P}$，又因 $\boldsymbol{B} \in \mathbf{C}^{n\times n}$ 是 **Hermitian** 矩阵，则 $\boldsymbol{Q}$ 也为 **Hermitian** 矩阵，故存在酉矩阵 $\boldsymbol{U}$，使得

$$\boldsymbol{U}^H\boldsymbol{Q}\boldsymbol{U} = \boldsymbol{\Lambda},$$

其中 $\boldsymbol{\Lambda} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \cdots, \boldsymbol{\lambda}_n)$ 为对角矩阵.

记 $\boldsymbol{T} = \boldsymbol{P}\boldsymbol{U}$，则有

$$\begin{aligned} \boldsymbol{T}^H\boldsymbol{A}\boldsymbol{T} &= \boldsymbol{U}^H\boldsymbol{P}^H\boldsymbol{A}\boldsymbol{P}\boldsymbol{U} = \boldsymbol{U}^H\boldsymbol{E}_n\boldsymbol{U} = \boldsymbol{E}_n, \\ \boldsymbol{T}^H\boldsymbol{B}\boldsymbol{T} &= \boldsymbol{U}^H\boldsymbol{P}^H\boldsymbol{B}\boldsymbol{P}\boldsymbol{U} = \boldsymbol{U}^H\boldsymbol{Q}\boldsymbol{U} = \boldsymbol{\Lambda}, \end{aligned}$$

故得证.

§4 矩阵的最大秩分解

程光辉

2020 年 4 月 4 日

定理 1 设 $\boldsymbol{A} \in \mathbf{C}_r^{m\times n}$，则存在矩阵 $\boldsymbol{B} \in \mathbf{C}_r^{m\times r}$，$\boldsymbol{D} \in \mathbf{C}_r^{r\times n}$，使得

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{D}.$$

证明： 因为 $\boldsymbol{A} \in \mathbf{C}_r^{m\times n}$，存在酉矩阵 $\boldsymbol{U} \in \mathbf{U}^{m\times m}, \boldsymbol{V} \in \mathbf{U}^{n\times n}$ 和 r 阶正线下三角矩阵 $\boldsymbol{L}$，使得

$$\begin{aligned}\boldsymbol{A} &= \boldsymbol{U}\begin{bmatrix}\boldsymbol{L} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{V}\\ &= \boldsymbol{U}\begin{bmatrix}\boldsymbol{L}\\ \boldsymbol{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{0}\end{bmatrix}\boldsymbol{V}\\ &= \boldsymbol{B}\boldsymbol{D},\end{aligned}$$

其中 $\boldsymbol{B} = \boldsymbol{U}\begin{bmatrix}\boldsymbol{L}\\ \boldsymbol{0}\end{bmatrix} \in \mathbf{C}_r^{m\times r}$，$\boldsymbol{D} = \begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{0}\end{bmatrix}\boldsymbol{V} \in \mathbf{C}_r^{r\times n}$，得证.

矩阵最大秩分解的步骤：

(1) 进行行初等变换，化为简化行阶梯型：

$$\tilde{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{0}\cdots\boldsymbol{0} & \boldsymbol{1} & \star\cdots\star & \boldsymbol{0} & \star\cdots\star & \boldsymbol{0} & \star\cdots\star\\ \boldsymbol{0}\cdots\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\cdots\boldsymbol{0} & \boldsymbol{1} & \star\cdots\star & \boldsymbol{0} & \star\cdots\star\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ \boldsymbol{0}\cdots\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\cdots\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\cdots\boldsymbol{0} & \boldsymbol{1} & \star\cdots\star\\ \boldsymbol{0}\cdots\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\cdots\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\cdots\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\cdots\boldsymbol{0}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ \boldsymbol{0}\cdots\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\cdots\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\cdots\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\cdots\boldsymbol{0} \end{bmatrix},$$

其中 $\star$ 表示不一定为 0 的元素，$\tilde{\boldsymbol{A}}$ 中非零首元所在的列分别为 $\boldsymbol{i_1}, \boldsymbol{i_2}, \cdots, \boldsymbol{i_r}$.

(2) 由 $\boldsymbol{A}$ 中第 $i_1, i_2, \cdots, i_r$ 列构成 $\boldsymbol{B} = (\boldsymbol{a}_{i_1}, \boldsymbol{a}_{i_2}, \cdots, \boldsymbol{a}_{i_r})$.

(3) 由 $\tilde{A}$ 的非零行构成 D.

思考上述算法的证明？还有哪些计算最大秩分解算法？

例 1 求矩阵 $A=\begin{bmatrix}1&3&2&1&4\\2&6&1&0&7\\3&9&3&1&11\end{bmatrix}$ 的最大秩分解.

解： (行初等变换) 对矩阵 A 进行行初等变换化为简化行阶梯型，有

$$A\rightarrow\begin{bmatrix}1&3&0&-\frac{1}{3}&\frac{10}{3}\\0&0&1&\frac{2}{3}&\frac{1}{3}\\0&0&0&0&0\end{bmatrix}=\tilde{A},$$

则

$$B=\begin{bmatrix}1&2\\2&1\\3&3\end{bmatrix},\quad D=\begin{bmatrix}1&3&0&-\frac{1}{3}&\frac{10}{3}\\0&0&1&\frac{2}{3}&\frac{1}{3}\end{bmatrix}.$$

(列初等变换) 对矩阵 A 进行列初等变换化为简化列阶梯型，有

$$A\rightarrow\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\1&1&0&0&0\end{bmatrix}=\tilde{A},$$

则

$$B=\begin{bmatrix}1&0\\0&1\\1&1\end{bmatrix},\quad D=\begin{bmatrix}1&3&2&1&4\\2&6&1&0&7\end{bmatrix}.$$

定理 2 设 $A\in C_r^{m\times n}$，且 $A=B_1D_1=B_2D_2$ 均是 A 的最大秩分解，则

(1) 存在 r 阶可逆矩阵 Q，使得

$$B_1=B_2Q,\quad D_1=Q^{-1}D_2.$$

(2) $D_1^H(D_1D_1^H)^{-1}(B_1^HB_1)^{-1}B_1^H=D_2^H(D_2D_2^H)^{-1}(B_2^HB_2)^{-1}B_2^H$.

注：

(1) $\mathrm{rank}(A)=\mathrm{rank}(A^HA)=\mathrm{rank}(AA^H)$.

(2) 若 $B\in C_r^{m\times r}$，则 $B^H\in C_r^{r\times m}$，$B^HB\in C_r^{r\times r}$，进而有

$$(B^HB)^{-1}(B^HB)=(B^HB)^{-1}B^HB=E_r.$$

(3) $\boldsymbol{D} \in \mathbf{C}_r^{r\times n}$，则 $\boldsymbol{D}^H \in \mathbf{C}_r^{n\times r}$，$\boldsymbol{D}\boldsymbol{D}^H \in \mathbf{C}_r^{r\times r}$，进而有

$$(\boldsymbol{D}\boldsymbol{D}^H)(\boldsymbol{D}\boldsymbol{D}^H)^{-1} = \boldsymbol{D}\boldsymbol{D}^H(\boldsymbol{D}\boldsymbol{D}^H)^{-1} = \boldsymbol{E}_r.$$

证明： (1) 因为 $\boldsymbol{B}_1\boldsymbol{D}_1 = \boldsymbol{B}_2\boldsymbol{D}_2$，则 $\boldsymbol{B}_1\boldsymbol{D}_1\boldsymbol{D}_1^H = \boldsymbol{B}_2\boldsymbol{D}_2\boldsymbol{D}_1^H$，进而得

$$\boldsymbol{B}_1 = \boldsymbol{B}_2\boldsymbol{D}_2\boldsymbol{D}_1^H\left(\boldsymbol{D}_1\boldsymbol{D}_1^H\right)^{-1} = \boldsymbol{B}_2\boldsymbol{Q}_1,$$

其中 $\boldsymbol{Q}_1 = \boldsymbol{D}_2\boldsymbol{D}_1^H\left(\boldsymbol{D}_1\boldsymbol{D}_1^H\right)^{-1}$.

同理可得 $\boldsymbol{D}_1 = (\boldsymbol{B}_1^H\boldsymbol{B}_1)^{-1}\boldsymbol{B}_1^H\boldsymbol{B}_2\boldsymbol{D}_2 = \boldsymbol{Q}_2\boldsymbol{D}_2$，其中 $\boldsymbol{Q}_2 = (\boldsymbol{B}_1^H\boldsymbol{B}_1)^{-1}\boldsymbol{B}_1^H\boldsymbol{B}_2$.

于是有

$$\boldsymbol{B}_1\boldsymbol{D}_1 = \boldsymbol{B}_2\boldsymbol{Q}_1\boldsymbol{Q}_2\boldsymbol{D}_2 = \boldsymbol{B}_2\boldsymbol{D}_2,$$

即

$$\boldsymbol{B}_2^H\boldsymbol{B}_2\boldsymbol{Q}_1\boldsymbol{Q}_2\boldsymbol{D}_2\boldsymbol{D}_2^H = \boldsymbol{B}_2^H\boldsymbol{B}_2\boldsymbol{D}_2\boldsymbol{D}_2^H,$$

因此，$\boldsymbol{Q}_1\boldsymbol{Q}_2 = \boldsymbol{E}_r$，记 $\boldsymbol{Q} = \boldsymbol{Q}_1$，则 $\boldsymbol{Q}_2 = \boldsymbol{Q}^{-1}$.

(2) 利用 (1)，有

$$\begin{aligned}
\boldsymbol{D}_1^H(\boldsymbol{D}_1\boldsymbol{D}_1^H)^{-1}(\boldsymbol{B}_1^H\boldsymbol{B}_1)^{-1}\boldsymbol{B}_1^H &= (\boldsymbol{Q}^{-1}\boldsymbol{D}_2)^H\left[\boldsymbol{Q}^{-1}\boldsymbol{D}_2(\boldsymbol{Q}^{-1}\boldsymbol{D}_2)^H\right]^{-1}\left[(\boldsymbol{B}_2\boldsymbol{Q})^H\boldsymbol{B}_2\boldsymbol{Q}\right]^{-1}(\boldsymbol{B}_2\boldsymbol{Q})^H \\
&= \boldsymbol{D}_2^H(\boldsymbol{Q}^{-1})^H\left[\boldsymbol{Q}^{-1}\boldsymbol{D}_2\boldsymbol{D}_2^H(\boldsymbol{Q}^{-1})^H\right]^{-1}\left[\boldsymbol{Q}^H\boldsymbol{B}_2^H\boldsymbol{B}_2\boldsymbol{Q}\right]^{-1}\boldsymbol{Q}^H\boldsymbol{B}_2^H \\
&= \boldsymbol{D}_2^H(\boldsymbol{Q}^{-1})^H\boldsymbol{Q}^H(\boldsymbol{D}_2\boldsymbol{D}_2^H)^{-1}\boldsymbol{Q}\boldsymbol{Q}^{-1}(\boldsymbol{B}_2^H\boldsymbol{B}_2)^{-1}(\boldsymbol{Q}^H)^{-1}\boldsymbol{Q}^H\boldsymbol{B}_2^H \\
&= \boldsymbol{D}_2^H(\boldsymbol{D}_2\boldsymbol{D}_2^H)^{-1}(\boldsymbol{B}_2^H\boldsymbol{B}_2)^{-1}\boldsymbol{B}_2^H,
\end{aligned}$$

得证.

§5 矩阵的奇异值分解

程光辉

2020 年 4 月 4 日

定理 1 设 $\boldsymbol{A} \in \mathbf{C}_r^{m\times n}$，则有

(1) $\mathbf{rank}(\boldsymbol{A}) = \mathbf{rank}(\boldsymbol{A^H A}) = \mathbf{rank}(\boldsymbol{AA^H})$

(2) $\boldsymbol{A^H A, AA^H}$ 的特征值均为非负实数

(3) $\boldsymbol{A^H A, AA^H}$ 的非零特征值相同.

证明： (1) 因为 $\mathbf{rank}(\boldsymbol{A^H A}) \le \mathbf{rank}(\boldsymbol{A})$，只需证明 $\mathbf{rank}(\boldsymbol{A^H A}) \ge \mathbf{rank}(\boldsymbol{A})$ 即可.

令 $\mathbf{rank}(\boldsymbol{A^H A}) = \boldsymbol{r}$，则 $\boldsymbol{A^H A x = 0}$ 的解空间为 $\boldsymbol{n-r}$ 维，记为 $\boldsymbol{X}$. 对 $\forall \boldsymbol{x_1} \in \boldsymbol{X}$，都有

$$\boldsymbol{x_1^H A^H A x_1 = (Ax_1)^H (Ax_1) = 0},$$

即

$$\boldsymbol{Ax_1 = 0}.$$

也就是说 $\boldsymbol{A^H A x = 0}$ 的解空间是 $\boldsymbol{Ax = 0}$ 的解空间的子空间，进而有

$$\mathbf{rank}(\boldsymbol{A^H A}) \ge \mathbf{rank}(\boldsymbol{A}),$$

得证.

(2) 设 $\boldsymbol{A^H A \alpha = \lambda \alpha}$，其中 $\boldsymbol{\lambda}$ 为任意特征值，$\boldsymbol{\alpha}$ 为对应的特征向量. 于是有

$$\boldsymbol{(A\alpha, A\alpha) = (\alpha, A^H A\alpha) = (\alpha, \lambda\alpha) = \lambda(\alpha, \alpha) \ge 0},$$

即 $\boldsymbol{\lambda \ge 0}$.

(3) 设 $\boldsymbol{A^H A}$ 的特征值为

$$\boldsymbol{\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0},$$

对应的特征向量分别为 $\boldsymbol{\alpha_1, \alpha_2, \cdots, \alpha_n}$.

因为 $A^HA\alpha_i = \lambda_i\alpha_i\ (\neq 0)$，$i = 1, 2, \cdots, r$，则

$$(AA^H)A\alpha_i = A(A^HA\alpha_i) = A(\lambda_i\alpha_i) = \lambda_i(A\alpha_i),$$

则 A^HA 的非零特征值 λ_i，$i = 1, 2, \cdots, r$ 也是 AA^H 的非零特征值，对应的特征向量为 $A\alpha_i$.

同理可证，AA^H 的非零特征值也是 A^HA 的非零特征值.

设 $y_1, \cdots, y_p$ 是 A^HA 的非零特征值 λ 对应的特征子空间 V_λ 的一组基，则有 $Ay_1, \cdots, Ay_p$ 是 AA^H 的非零特征值 λ 对应的特征向量.

令

$$k_1Ay_1 + k_2Ay_2 + \cdots + k_pAy_p = 0,$$

左乘 A^H，有

$$k_1A^HAy_1 + k_2A^HAy_2 + \cdots + k_pA^HAy_p = \lambda(k_1y_1 + k_2y_2 + \cdots + k_py_p) = 0,$$

即只有 $k_1 = k_2 = \cdots = k_p = 0$，则 $Ay_1, \cdots, Ay_p$ 线性无关. 也就是说，A^HA 的特征子空间 V_λ 的维数小于等于 AA^H 的特征子空间 V_λ 的维数.

同理可证，AA^H 的特征子空间 V_λ 的维数小于等于 A^HA 的特征子空间 V_λ 的维数.

综上，得证.

定义 1 设 $A \in \mathbf{C}_r^{m\times n}$，$A^HA$ 的特征值为

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0,$$

则称 $\sigma_i = \sqrt{\lambda_i} \quad (i = 1, 2, \cdots, r)$ 为 A 的正奇异值.

定义 2 设 $A, B \in \mathbf{C}^{m\times n}$，如果存在酉矩阵 $U \in \mathbf{U}^{m\times m}$ 和 $V \in \mathbf{U}^{n\times n}$，使得

$$A = UBV,$$

则称 A 与 B 酉等价.

定理 2 若 A 与 B 酉等价，则 A 与 B 有相同的正奇异值.

证明: 因为 A 与 B 酉等价，则存在酉矩阵 $U \in \mathbf{U}^{m\times m}$ 和 $V \in \mathbf{U}^{n\times n}$，使得 $A = UBV$，进而有

$$AA^H = UBV(UBV)^H = UBVV^HB^HU^H = UBB^HU^H,$$

所以 AA^H 和 BB^H 相似，有相同的特征值，即 A 与 B 有相同的正奇异值.

定理 3 设 $A \in \mathbf{C}_r^{m\times n}$，$\sigma_1, \sigma_2, \cdots, \sigma_r$ 是矩阵 A 的 r 个正奇异值，则存在酉矩阵 $U \in \mathbf{U}^{m\times m}$ 和 $V \in \mathbf{U}^{n\times n}$，使得

$$A = U\begin{bmatrix} D & O \\ O & O \end{bmatrix}V^H,$$

其中 $D = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_r)$.

证明：因为 A^HA 为 Hermitian 半正定矩阵，故特征值为 $\sigma_1^2, \sigma_2^2, \cdots, \sigma_r^2, 0, \cdots, 0$. 令 $v_1, v_2, \cdots, v_n$ 是对应的单位正交特征向量组，而且

$$V_1 = [v_1, v_2, \cdots, v_r]$$
$$V_2 = [v_{r+1}, v_{r+2}, \cdots, v_n].$$

因此，有 $A^HAV_1 = V_1D^2$，进而

$$D^{-1}V_1^HA^HAV_1D^{-1} = U_1^HU_1 = E_r,$$

其中 $U_1 = AV_1D^{-1} \in \mathrm{U}_r^{m\times r}$，即 $D = U_1^HAV_1$，$U_1D = AV_1$.

另有 $A^HAV_2 = V_2O$，使得 $V_2^HA^HAV_2 = O$，因此 $AV_2 = O$.

选择合适 $U_2 \in \mathrm{U}_{(m-r)}^{m\times(m-r)}$，使得 $U = [U_1, U_2]$ 为酉矩阵. 于是

$$U^HAV = \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix} A[V_1, V_2] = \begin{bmatrix} U_1^HAV_1 & U_1^HAV_2 \\ U_2^HAV_1 & U_2^HAV_2 \end{bmatrix} = \begin{bmatrix} D & O \\ U_2^HU_1D & O \end{bmatrix} = \begin{bmatrix} D & O \\ O & O \end{bmatrix},$$

从而有

$$A = U\begin{bmatrix} D & O \\ O & O \end{bmatrix}V^H.$$

推论 1 设 $A \in \mathrm{C}_n^{n\times n}$, $\sigma_1, \sigma_2, \cdots, \sigma_n$ 是矩阵 A 的 n 个正奇异值, 则存在酉矩阵 $U, V \in \mathrm{U}^{n\times n}$, 使得

$$A = UDV^H,$$

其中 $D = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_n)$.

例 1 求矩阵 $A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$ 的奇异值分解.

解：求 A^HA 的特征值和特征向量，

$$A^HA = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

则特征值为 $\lambda_1 = 5, \lambda_2 = \lambda_3 = 0$. 对应的特征向量为

$$\alpha_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\ \alpha_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},\ \alpha_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

因此，有

$$V_1=\begin{bmatrix}1\\0\\0\end{bmatrix},\ V_2=\begin{bmatrix}0&0\\1&0\\0&1\end{bmatrix},\ V=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}.$$

构造酉矩阵 U，有

$$U_1=AV_1D^{-1}=\begin{bmatrix}\frac{1}{\sqrt5}\\\frac{2}{\sqrt5}\end{bmatrix},\ U_2=\begin{bmatrix}-\frac{2}{\sqrt5}\\\frac{1}{\sqrt5}\end{bmatrix},\ U=\begin{bmatrix}\frac{1}{\sqrt5}&-\frac{2}{\sqrt5}\\\frac{2}{\sqrt5}&\frac{1}{\sqrt5}\end{bmatrix}.$$

综上，有

$$A=\begin{bmatrix}\frac{1}{\sqrt5}&-\frac{2}{\sqrt5}\\\frac{2}{\sqrt5}&\frac{1}{\sqrt5}\end{bmatrix}\begin{bmatrix}\sqrt5&0&0\\0&0&0\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}.$$

引理 1 设 $A\in C_r^{n\times n}$ 是半正定矩阵，且 $k\geq 1$ 是给定的整数，则存在唯一的半正定 Hermitian 矩阵 B 使得 $B^k=A$. 同时还有

(1) $BA=AB$,

(2) $\mathrm{rank}(A)=\mathrm{rank}(B)$, 进而若 A 正定，则 B 也正定，

(3) 若 A 是实矩阵，则 B 也实矩阵.

定理 4 设 $A\in C_n^{n\times n}$, 则必存在酉矩阵 $U\in U^{n\times n}$ 与两个正定 Hermitian 矩阵 H_1, H_2, 使得

$$A=H_1U=UH_2,$$

且这种分解式是唯一的.

证明： (存在性) 因为 $A\in C_n^{n\times n}$，所以 A^HA 为正定 Hermitian 矩阵，则矩阵 A 的奇异值均为正数，其奇异值分解为

$$A=U_1DV_1^H=U_1DU_1^HU_1V_1^H=H_1U,$$

其中 $H_1=U_1DU_1^H$ 为 Hermitian 正定矩阵，$U=U_1V_1^H$ 为酉矩阵.

(唯一性) 令 $A=H_{11}U_1=H_{12}U_2$, 其中 H_{11}, H_{12} 为正定 Hermitian 矩阵, U_1, U_2 为酉矩阵. 因此，有

$$H_{12}^{-1}H_{11}=U_2U_1^H$$
$$H_{11}^{-1}H_{12}=U_1U_2^H,$$

上面两个等式的右侧矩阵互为共轭转置，进而

$$H_{12}^{-1}H_{11}=\left(H_{11}^{-1}H_{12}\right)^H=H_{12}^H\left(H_{11}^{-1}\right)^H=H_{12}H_{11}^{-1},$$

即

$$\boldsymbol{H}_{11}^2 = \boldsymbol{H}_{12}^2,$$

再由引理 1，知

$$\boldsymbol{H}_{11} = \boldsymbol{H}_{12}, \quad \boldsymbol{U}_1 = \boldsymbol{U}_2,$$

得证.

推论 2 设 $\boldsymbol{A} \in \mathbf{R}_n^{n\times n}$，则必存在唯一的正交矩阵 $\boldsymbol{Q}$ 与两个实对称正定矩阵 $\boldsymbol{H}_1$，$\boldsymbol{H}_2$，使得

$$\boldsymbol{A} = \boldsymbol{H}_1\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{H}_2.$$

定理 5 设 $\boldsymbol{A} \in \mathbf{C}_r^{n\times n}$，则必存在酉矩阵 $\boldsymbol{U} \in \mathbf{U}^{n\times n}$ 与两个半正定 **Hermitian** 矩阵 $\boldsymbol{H}_1$，$\boldsymbol{H}_2$，使得

$$\boldsymbol{A} = \boldsymbol{H}_1\boldsymbol{U} = \boldsymbol{U}\boldsymbol{H}_2.$$

证明： 因为 $\boldsymbol{A} \in \mathbf{C}_r^{n\times n}$，则由奇异值分解可得

$$\boldsymbol{A} = \boldsymbol{U}_1\begin{bmatrix}\boldsymbol{D} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\boldsymbol{V}_1^H = \boldsymbol{U}_1\begin{bmatrix}\boldsymbol{D} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\boldsymbol{U}_1^H\boldsymbol{U}_1\boldsymbol{V}_1^H = \boldsymbol{H}_1\boldsymbol{U},$$

其中 $\boldsymbol{H}_1 = \boldsymbol{U}_1\begin{bmatrix}\boldsymbol{D} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\boldsymbol{U}_1^H$ 为半正定 **Hermitian** 矩阵，$\boldsymbol{U} = \boldsymbol{U}_1\boldsymbol{V}_1^H$ 为酉矩阵.

§1 特征值的估计

程光辉

2020 年 4 月 28 日

1　特征值界的估计

定理 1. *(***Schur** *不等式)* 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{n\times n}$ 的特征值为 $\boldsymbol{\lambda_1,\lambda_2,\cdots,\lambda_n}$，则

$$\sum_{i=1}^{n}|\boldsymbol{\lambda}_i|^2\leq\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2=\|\boldsymbol{A}\|_F^2,$$

且等号成立当且仅当 $\boldsymbol{A}$ 为正规矩阵.

证明： 由 **Schur** 分解知，存在上三角矩阵 $\boldsymbol{R}=(r_{ij})$ 和酉矩阵 $\boldsymbol{U}$，使得 $\boldsymbol{A}=\boldsymbol{URU}^H$. 又由 **Frobenius** 范数的酉不变性，知

$$\sum_{i=1}^{n}|\boldsymbol{\lambda}_i|^2=\sum_{i=1}^{n}|r_{ii}|^2\leq\sum_{i=1}^{n}|r_{ii}|^2+\sum_{i=1}^{n}\sum_{j=i+1}^{n}|r_{ij}|^2=\|\boldsymbol{R}\|_F^2=\|\boldsymbol{A}\|_F^2,$$

证毕.

令

$$\boldsymbol{B}=(b_{ij})=\frac{1}{2}\left(\boldsymbol{A}^H+\boldsymbol{A}\right),\quad \boldsymbol{C}=(c_{ij})=\frac{1}{2}\left(\boldsymbol{A}-\boldsymbol{A}^H\right),$$

$\boldsymbol{A,B,C}$ 的特征值分别为 $\{\boldsymbol{\lambda_1,\lambda_2,\cdots,\lambda_n}\},\{\boldsymbol{\mu_1,\mu_2,\cdots,\mu_n}\},\{i\gamma_1,i\gamma_2,\cdots,i\gamma_n\}$，且满足

$$\begin{aligned}|\boldsymbol{\lambda}_1|\geq|\boldsymbol{\lambda}_2|\geq\cdots\geq|\boldsymbol{\lambda}_n|,\\ \boldsymbol{\mu}_1\geq\boldsymbol{\mu}_2\geq\cdots\geq\boldsymbol{\mu}_n,\\ \gamma_1\geq\gamma_2\geq\cdots\geq\gamma_n.\end{aligned}$$

定理 2. *(*Hirsch*)* 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$ 的特征值为 $\boldsymbol{\lambda_1,\lambda_2,\cdots,\lambda_n}$，则

(1) $|\boldsymbol{\lambda}_i|\leq n\max\limits_{i,j}\{|a_{ij}|\}$,

(2) $|\mathbf{Re}\boldsymbol{\lambda}_i|\leq n\max\limits_{i,j}\{|b_{ij}|\}$,

(3) $|\mathrm{Im}\,\boldsymbol{\lambda}_i| \leq n \max\limits_{i,j}\{|c_{ij}|\}$.

证明：(1) 由

$$|\boldsymbol{\lambda}_i|^2 \leq \sum_{i=1}^{n} |\boldsymbol{\lambda}_i|^2 \leq \sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^2 \leq n^2 \max_{i,j}\left\{|a_{ij}|^2\right\},$$

得 $|\boldsymbol{\lambda}_i| \leq n \max\limits_{i,j}\{|a_{ij}|\}$.

(2)(3) 因为 $\boldsymbol{A} \in \mathrm{C}^{n\times n}$，存在上三角矩阵 $\boldsymbol{R}$ 和酉矩阵 $\boldsymbol{U}$，使得 $\boldsymbol{U}^{\boldsymbol{H}}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{R}$，$\boldsymbol{U}^{\boldsymbol{H}}\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{U} = \boldsymbol{R}^{\boldsymbol{H}}$，那么

$$\begin{cases} \boldsymbol{U}^H\boldsymbol{B}\boldsymbol{U} = \frac{1}{2}\boldsymbol{U}^H\left(\boldsymbol{A}^H + \boldsymbol{A}\right)\boldsymbol{U} = \frac{1}{2}\left(\boldsymbol{R}^H + \boldsymbol{R}\right) \\ \boldsymbol{U}^H\boldsymbol{C}\boldsymbol{U} = \frac{1}{2}\boldsymbol{U}^H\left(\boldsymbol{A} - \boldsymbol{A}^H\right)\boldsymbol{U} = \frac{1}{2}\left(\boldsymbol{R} - \boldsymbol{R}^H\right) \end{cases},$$

由 **Frobenius** 范数酉的不变性，可得

$$\begin{cases} \sum\limits_{i=1}^{n}\left|\frac{\lambda_i+\bar{\lambda}_i}{2}\right|^2 + \sum\limits_{j=1}^{n}\sum\limits_{i=1}^{j-1}\frac{|r_{ij}|^2}{2} = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}|b_{ij}|^2 \\ \sum\limits_{i=1}^{n}\left|\frac{\lambda_i-\bar{\lambda}_i}{2}\right|^2 + \sum\limits_{j=1}^{n}\sum\limits_{i=1}^{j-1}\frac{|r_{ij}|^2}{2} = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}|c_{ij}|^2 \end{cases},$$

所以有

$$\begin{cases} |\mathrm{Re}\,\boldsymbol{\lambda}_i|^2 \leq \sum\limits_{i=1}^{n}|\mathrm{Re}\,\boldsymbol{\lambda}_i|^2 \leq \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}|b_{ij}|^2 \leq n^2\max\limits_{i,j}\{|b_{ij}|^2\} \\ |\mathrm{Im}\,\boldsymbol{\lambda}_i|^2 \leq \sum\limits_{i=1}^{n}|\mathrm{Im}\,\boldsymbol{\lambda}_i|^2 \leq \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}|c_{ij}|^2 \leq n^2\max\limits_{i,j}\{|c_{ij}|^2\} \end{cases},$$

因此，

$$\begin{cases} |\,\mathrm{Re}\,\boldsymbol{\lambda}_i| \leq n\max\limits_{i,j}\{|b_{ij}|\} \\ |\,\mathrm{Im}\,\boldsymbol{\lambda}_i| \leq n\max\limits_{i,j}\{|c_{ij}|\} \end{cases}$$

证毕.

另证：$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\lambda}\boldsymbol{x}$，$\|\boldsymbol{x}\|_2 = 1$，则 $\mathrm{Re}\,\boldsymbol{\lambda} = \boldsymbol{x}^H\boldsymbol{B}\boldsymbol{x}$，由于 $\boldsymbol{B}$ 为 **Hermitian** 矩阵，酉相似于对角矩阵，即有

$$|\,\mathrm{Re}\,\boldsymbol{\lambda}| = |\boldsymbol{x}^H\boldsymbol{B}\boldsymbol{x}| \leq \max_i\{|\mu_i|\} \leq n\max_{i,j}\{|b_{ij}|\}.$$

定理 3. *(Bendixson)* 设 $\boldsymbol{A} \in \mathrm{R}^{n\times n}$，则 $\boldsymbol{A}$ 的任一特征值 $\boldsymbol{\lambda}_i$ 满足

$$|\mathrm{Im}\,\boldsymbol{\lambda}_i| \leq \sqrt{\frac{n(n-1)}{2}}\max_{i,j}\{|c_{ij}|\}.$$

证明：因为

$$\sum_{i=1}^{n}|\mathrm{Im}\,\boldsymbol{\lambda}_i|^2 \leq \sum_{i=1}^{n}\sum_{j=1}^{n}|c_{ij}|^2 = \sum_{\substack{i,j\\ i\neq j}}|c_{ij}|^2 \leq n(n-1)\max_{i,j}\{|c_{ij}|^2\}.$$

令互为共轭的特征值一共 s 对，再根据实矩阵复特征值出现一定是成对 (互为共轭)，有

$$2\left|\mathrm{Im}\,\lambda_i\right|^2 \leq 2\sum_{i=1}^{s}\left|\mathrm{Im}\,\lambda_i\right|^2 = \sum_{i=1}^{n}\left|\mathrm{Im}\,\lambda_i\right|^2 \leq n(n-1)\max_{i,j}\{|c_{ij}|^2\},$$

整理即可得证.

定理 4. 设 $A \in \mathbf{C}^{n\times n}, B, C, \lambda_i, \mu_i, \gamma_i$ 定义同上, 则

$$\mu_n \leq \mathrm{Re}\lambda_i \leq \mu_1, \gamma_n \leq \mathrm{Im}\lambda_i \leq \gamma_1.$$

证明： 由 B 为正规矩阵，存在酉矩阵 U 使得

$$U^H B U = \mathrm{diag}\left(\mu_1, \mu_2, \cdots, \mu_n\right) = D.$$

因为 $Ax = \lambda_i x\,(\|x\|_2 = 1)$，$(x, Ax) = (x, \lambda_i x) = \lambda_i(x, x) = \lambda_i$，有 $x^H A x = \lambda_i$ 和 $x^H A^H x = \bar{\lambda}_i$.

于是有

$$\mathrm{Re}\,\lambda_i = \left(x, \frac{A^H + A}{2}x\right) = (x, Bx) = \left(x, UDU^H x\right)$$

$$= x^H UDU^H x = y^H D y = \sum_{i=1}^{n}\mu_i\left|y_i\right|^2,$$

其中 $y = U^H x$，$\|y\|_2 = \|x\|_2 = 1$. 进而可得

$$\sum_{i=1}^{n}\mu_n\left|y_i\right|^2 \leq \mathrm{Re}\,\lambda_i \leq \sum_{i=1}^{n}\mu_1\left|y_i\right|^2,$$

故

$$\mu_n \leq \mathrm{Re}\lambda_i \leq \mu_1.$$

同理，可证 $\gamma_n \leq \mathrm{Im}\lambda_i \leq \gamma_1$.

定理 5. *(Browne)* 设 $A \in \mathbf{C}^{n\times n}$ 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$，奇异值为 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$，则

$$\sigma_n \leq |\lambda_i| \leq \sigma_1, \quad (i = 1, 2, \cdots, n).$$

证明： 因 $A^H A$ 为 Hermitian 矩阵，存在酉矩阵 U 使得

$$U^H A^H A U = \mathrm{diag}\left(\sigma_1^2, \sigma_2^2, \cdots, \sigma_n^2\right) = D.$$

令 $Ax = \lambda_i x$，$(\|x\|_2 = 1)$，则

$$x^H A^H = \bar{\lambda}_i x^H,$$

进而
$$x^H A^H A x = |\lambda_i|^2 x^H x = |\lambda_i|^2 .$$

又因为
$$|\lambda_i|^2 = x^H A^H A x = x^H U D U^H x = \sum_{i=1}^{n} \sigma_i^2 |y_i|^2,$$
其中 $y = U^H x = (y_1, \cdots, y_n)^T$，由范数的酉不变性知 $\|y\|_2 = \|x\|_2 = 1$，因此
$$\sigma_n^2 = \sigma_n^2 \sum_{i=1}^{n} |y_i|^2 \leq |\lambda_i|^2 \leq \sigma_1^2 \sum_{i=1}^{n} |y_i|^2 = \sigma_1^2,$$
两边开方即得证.

定理 6. *(Hadamard 不等式)* 设 $A = (a_1, \cdots, a_n) \in \mathbf{C}^{n\times n}$, 则
$$\prod_{i=1}^{n} |\lambda_i(A)| = |\det A| \leq \left[\prod_{j=1}^{n} \left(\sum_{i=1}^{n} |a_{ij}|^2 \right) \right]^{1/2} = \prod_{j=1}^{n} \|a_j\|_2,$$
且等号成立当且仅当 A 的某一列全为 O, 或 A 的列向量彼此正交.

证明： 1) 设 $A = (a_1, a_2, \cdots, a_n)$，若 $a_1, a_2, \cdots, a_n$ 线性相关，则 $|\det A| = 0$，则成立.

2) 若 $a_1, a_2, \cdots, a_n$ 线性无关，对其正交化有
$$\left.\begin{aligned} & a_1 = b_1 \\ & a_2 = b_2 + p_{12} b_1 \\ & a_3 = b_3 + p_{13} b_1 + p_{23} b_2 \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ & a_n = b_n + p_{1n} b_1 + \cdots + p_{n-1,n} b_{n-1} \end{aligned}\right\},$$
即
$$A = (b_1, b_2, \cdots, b_n) \begin{bmatrix} 1 & p_{12} & \cdots & p_{1n} \\ 0 & 1 & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$
两边取行列式有
$$\det A = \det (b_1, b_2, \cdots, b_n) = \det B,$$
进而
$$\begin{aligned} \|a_i\|_2^2 &= \|b_i + p_{1i} b_1 + \cdots + p_{i-1,i} b_{i-1}\|_2^2 \\ &= \|b_i\|_2^2 + |p_{i1}| \, \|b_1\|_2^2 + \cdots + |p_{i,i-1}| \, \||b_{i-1}\|_2^2 \\ &\geq \|b_i\|_2^2 . \end{aligned}$$

又

$$|\det B|^2 = \det B^H \cdot \det B = \det B^H B = \prod_{j=1}^{n} \|b_j\|_2^2 \le \left(\prod_{j=1}^{n} \|a_j\|_2 \right)^2,$$

故

$$\prod_{i=1}^{n} |\lambda_i(A)| = |\det A| \le \left[\prod_{j=1}^{n} \left(\sum_{i=1}^{n} |a_{ij}|^2 \right) \right]^{1/2} = \prod_{j=1}^{n} \|a_j\|_2.$$

另解: 若 A 满秩, 则存在酉矩阵 U 和正线上三角矩阵 R, 使得 $A = UR$. 因 Frobenius 范数的酉不变性，知 $\|a_i\|_2 = \|r_i\|_2$, $i = 1, 2, \cdots, n$. 又因

$$|\det A| = |\det UR| = |\det U| \cdot |\det R| = |\det R| \le \prod_{j=1}^{n} \|r_j\|_2 = \prod_{j=1}^{n} \|a_j\|_2,$$

故

$$\prod_{i=1}^{n} |\lambda_i(A)| = |\det A| \le \left[\prod_{j=1}^{n} \left(\sum_{i=1}^{n} |a_{ij}|^2 \right) \right]^{1/2} = \prod_{j=1}^{n} \|a_j\|_2.$$

§2 圆盘定理

程光辉

2020 年 4 月 28 日

定义 1. 设 $\boldsymbol{A}=(a_{ij})\in \mathbf{C}^{n\times n}$，则称

$$\boldsymbol{S}_i=\left\{z\in C:|z-a_{ii}|\le R_i=\sum_{j\ne i}|a_{ij}|\right\}$$

为行盖尔圆盘;

$$\boldsymbol{G}_i=\left\{z\in C:|z-a_{ii}|\le C_i=\sum_{j\ne i}|a_{ji}|\right\}$$

为列盖尔圆盘.

定理 1. *(1931, 圆盘定理 1)* 设 $\boldsymbol{A}\in \mathbf{C}^{n\times n}$，则 $\boldsymbol{A}$ 的任一特征值

$$\lambda\in S=\bigcup_{j=1}^{n}S_j.$$

证明： 因为 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$, $\left(\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^T\ne\mathbf{0}\right)$，则

$$\sum_{j=1}^{n}a_{ij}x_j=\lambda x_i\ (i=1,2,\cdots,n).$$

令 $|x_k|=\max\{|x_1|,\cdots,|x_n|\}>0$，那么

$$\sum_{j=1}^{n}a_{kj}x_j=\lambda x_k,$$

移项得

$$x_k(\lambda-a_{kk})=\sum_{j\ne k}a_{kj}x_j,$$

两边取模放缩有

$$|x_k||\lambda-a_{kk}|=\left|\sum_{j\ne k}a_{kj}x_j\right|\le\sum_{j\ne k}|a_{kj}|\,|x_j|\le|x_k|\sum_{j\ne k}|a_{kj}|$$

故

$$|\lambda - a_{kk}| \leq R_k,$$

由于 λ 的任意性，得证.

例 1. 估计矩阵

$$A = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{3}{2} & i & 0 \\ 0 & -\frac{i}{2} & 5 & -\frac{i}{2} \\ -1 & 0 & 0 & 5i \end{bmatrix}$$

的特征值的分布范围.

解:

$$\begin{aligned} S_1 &: |z-1| \leq 1; \\ S_2 &: \left|z-\frac{3}{2}\right| \leq \frac{3}{2}; \\ S_3 &: |z-5| \leq 1; \\ S_4 &: |z-5i| \leq 1. \end{aligned}$$

推论 1. 设 $A \in \mathbf{C}^{n\times n}$，则 A 的任一特征值

$$\lambda \in \bigcup_{i=1}^{n} G_i.$$

定理 2. *(圆盘定理 2)* 设 n 阶方阵 $A=(a_{ij})$ 的 n 个盖尔圆盘中有 k 个圆盘的并形成一个连通区域 G，且它与余下的 $n-k$ 个圆盘都不相交，则在该区域 G 中恰好有 A 的 k 个特征值.

证明: 把矩阵 A 分解为对角部分和非对角部分的和，即

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} + \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & 0 \end{bmatrix} =: D + B.$$

令

$$A_\varepsilon = D + \varepsilon B, \quad \varepsilon \in [0,1],$$

则

$$R_i\left(A_\varepsilon\right) = R_i(\varepsilon B) = \varepsilon R_i(A).$$

设

$$\boldsymbol{G}_k=\bigcup_{i=1}^{k}\{z\in C:|z-a_{ii}|\leq R_i(A)\},$$
$$\boldsymbol{G}_k(\varepsilon)=\bigcup_{i=1}^{k}\{z\in C:|z-a_{ii}|\leq R_i(A_\varepsilon)=\varepsilon R_i(A)\},$$

故 $G_k\equiv G_k(1)$.

根据特征多项式知，矩阵特征值是关于矩阵元素的连续函数，也就是说 A_ε 的特征值是关于 ε 的连续函数，即特征值随着 ε 变化的曲线是不间断的. 当 ε 从 0 到 1 变化时，特征值 $\lambda(A_\varepsilon)$ 从矩阵 A 的对角元素为起点到 A 的对应特征值为终点变化. 因为特征值不会跑到区域外或者消失，也不会有外面的特征值进入区域 G，因此，k 个圆盘的并形成一个连通区域 G 中恰好有 k 个特征值.

推论 2. 设 $A\in\mathbf{C}^{n\times n}$，$\lambda$ 是 A 的任一特征值，则

$$\lambda\in\left(\bigcup_{i=1}^{n}S_i\right)\cap\left(\bigcup_{j=1}^{n}G_j\right).$$

推论 3. 设 n 阶方阵 A 的 n 个盖尔圆盘两两互不相交, 则 A 相似于对角阵.

推论 4. 设 n 阶实阵 A 的 n 个盖尔圆盘两两互不相交, 则 A 特征值全为实数.

令 $D=\operatorname{diag}(p_1,p_2,\cdots,p_n),(p_i>0)$，则

$$D^{-1}AD=\begin{bmatrix}a_{11}&\frac{p_2}{p_1}a_{12}&\cdots&\frac{p_n}{p_1}a_{1n}\\\frac{p_1}{p_2}a_{21}&a_{22}&\cdots&\frac{p_n}{p_2}a_{2n}\\\cdots&\cdots&\cdots&\cdots\\\frac{p_1}{p_n}a_{n1}&\frac{p_2}{p_n}a_{n2}&\cdots&a_{nn}\end{bmatrix},$$

进而

$$r_i=\frac{1}{p_i}\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|\,p_j,\quad Q_i=\{z\in C:|z-a_{ii}|\leq r_i\},$$
$$t_j=p_j\sum_{\substack{i=1\\i\neq j}}^{n}\frac{|a_{ij}|}{p_i},\quad P_j=\{z\in C:|z-a_{jj}|\leq t_j\}.$$

定理 3. 设 $A\in\mathbf{C}^{n\times n}$，$\lambda$ 是 A 的任一特征值，则

$$\lambda\in\left(\bigcup_{i=1}^{n}Q_i\right)\cap\left(\bigcup_{j=1}^{n}P_j\right).$$

例 2. 估计矩阵 $\boldsymbol{A}=\begin{bmatrix}0.9 & 0.01 & 0.12\\ 0.01 & 0.8 & 0.13\\ 0.01 & 0.02 & 0.4\end{bmatrix}$ 的特征值范围.

解： 行盖尔圆为

$$
\begin{aligned}
S_1 &:|z-0.9|\le 0.13;\\
S_2 &:|z-0.8|\le 0.14;\\
S_3 &:|z-0.4|\le 0.03.
\end{aligned}
$$

先令 $D=\mathrm{diag}(1,1,0.1)$，对 A 进行相似变换有 $D^{-1}AD$，其行盖尔圆为

$$
\begin{aligned}
S_1 &:|z-0.9|\le 0.022\\
S_2 &:|z-0.8|\le 0.023\\
S_3 &:|z-0.4|\le 0.3.
\end{aligned}
$$

两者结合得更好估计.

定义 2. 设 $A\in\mathrm{C}^{n\times n}$，若 A 行对角占优，则

$$|a_{ii}|\ge R_i=\sum_{\substack{j=1\\ j\ne i}}^{n}|a_{ij}|,\quad (i=1,2,\cdots,n);$$

若 A 列对角占优, 则

$$|a_{ii}|\ge C_i=\sum_{\substack{j=1\\ j\ne i}}^{n}|a_{ji}|,\quad (i=1,2,\cdots,n);$$

若 A 行严格对角占优, 则

$$|a_{ii}|> R_i=\sum_{\substack{j=1\\ j\ne i}}^{n}|a_{ij}|,\quad (i=1,2,\cdots,n);$$

若 A 列严格对角占优, 则

$$|a_{ii}|> C_i=\sum_{\substack{j=1\\ j\ne i}}^{n}|a_{ji}|,\quad (i=1,2,\cdots,n).$$

定理 4. 设 $A\in\mathrm{C}^{n\times n}$ 行 (或列) 严格对角占优，则

(1) A 可逆, 且 $\lambda_i\in\bigcup\limits_{i=1}^{n}S_i$, $\quad(S_i=\{z\in C:|z-a_{ii}|\le|a_{ii}|\})$;

(2) 若 A 的所有主对角元都是正数, 则 A 的特征值都有正实部;

(3) 若 $\boldsymbol{A}$ 为 $\boldsymbol{Hermitian}$ 矩阵，且所有主对角元都为正数，则 $\boldsymbol{A}$ 的特征值都为正数.

证明：(1) 因为 $\boldsymbol{A}$ 行严格对角占优，则有 $\boldsymbol{R_i = \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| < |a_{ii}|}$，则

$$\boldsymbol{\lambda_i \in \bigcup_{i=1}^{n} S_i, \quad (S_i = \{z \in C : |z - a_{ii}| < |a_{ii}|\})\,;}$$

显然 $\boldsymbol{0 \notin S_i}$，所以 $\boldsymbol{0}$ 不是其特征值，即可逆.

(2) 因为 $\boldsymbol{a_{ii} > 0, |\lambda - a_{ii}| < |a_{ii}|}$，则 $\boldsymbol{A}$ 的特征值都有正实部.

(3)$\boldsymbol{A^H = A}$，则 $\boldsymbol{A}$ 得特征值都是实数，所以 $\boldsymbol{A}$ 的特征值都有正数.

§4 Hermitian 矩阵特征值的变分特征

程光辉

2020 年 4 月 28 日

定义 1. *设 $A \in \mathbf{C}^{n\times n}$ 为 Hermitian 矩阵，$x \in \mathbf{C}^n$，称*

$$R(x) = \frac{x^H A x}{x^H x}, \quad x \neq 0$$

为 A 的 Rayleigh 商.

若 $\|x\| = 1$，则 **Rayleigh** 商变为

$$\mathrm{R}(x) = x^H A x.$$

为了下文方便，给出如下结论和记号：

若 $A \in \mathbf{C}^{n\times n}$ 为 **Hermitian** 矩阵，则存在酉矩阵 $U = (u_1, u_2, \cdots, u_n)$ 使得 $A = U\Lambda U^H$，其中 $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$，且 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

引理 1. *若 $W = span\{u_r, \cdots, u_s\}$，其中 $1 \leq r \leq s \leq n$，则对 $\forall x \in W$，$\|x\| = 1$，有*

$$\lambda_s \leq x^H A x \leq \lambda_r.$$

证明： 对 $\forall x \in W$，$\|x\| = 1$，有

$$x = \sum_{i=r}^{s} k_i u_i, \quad \sum_{i=r}^{s} |k_i|^2 = 1.$$

进而

$$x^H A x = x^H \sum_{i=r}^{s} k_i A u_i = x^H \sum_{i=r}^{s} k_i \lambda_i u_i = \sum_{i=r}^{s} |k_i|^2 \lambda_i.$$

又因 $\lambda_s \leq \lambda_i \leq \lambda_r$，$i = r, \cdots, s$，则有

$$\lambda_s = \lambda_s \sum_{i=r}^{s} |k_i|^2 \leq x^H A x = \sum_{i=r}^{s} |k_i|^2 \lambda_i \leq \lambda_r \sum_{i=r}^{s} |k_i|^2 = \lambda_r,$$

得证.

定理 1. *(Rayleigh -Ritz)* 设 $A \in \mathbf{C}^{n\times n}$ 为 *Hermitian* 矩阵，则

(1) $\lambda_n x^H x \le x^H A x \le \lambda_1 x^H x, \quad (\forall x \in \mathbf{C}^n)$，

(2) $\lambda_{\max} = \lambda_1 = \max\limits_{x\neq 0} R(x) = \max\limits_{x^H x=1} x^H A x$,

(3) $\lambda_{\min} = \lambda_n = \min\limits_{x\neq 0} R(x) = \min\limits_{x^H x=1} x^H A x$.

证明：(1) 直接由引理 1，即可得证.

(2) 取 x 为 λ_1 对应的特征向量，即可.

(3) 取 x 为 λ_n 对应的特征向量，即可.

定理 2. *(Courant -Fischer)* 设 $A \in \mathbf{C}^{n\times n}$ 为 *Hermitian* 矩阵，特征值为 $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_n$，$k$ 为给定的正整数，且 $1 \le k \le n$，则

$$\max_{V, dim(V)=k} \quad \min_{x\in V, \|x\|=1} R(x) = \lambda_k,$$

$$\min_{V, dim(V)=n-k+1} \quad \max_{x\in V, \|x\|=1} R(x) = \lambda_k.$$

证明：设 A 的特征值 λ_i 对应的单位特征向量为 u_i. 考虑如下子空间

$$T = \mathrm{span}\{u_k, \cdots, u_n\},$$

则 $\dim(T) = n-(k-1) = n-k+1$.

设 V 是 $\mathbf{C}^n$ 上任意的子空间，且 $\dim(V) = k$. 因为

$$\dim(T) + \dim(V) = n-k+1+k = n+1,$$

则有 $\dim(T\cap V) \ge 1$. 对 $\forall x \in T\cap V$，且 $\|x\| = 1$，因 $x \in T$，由引理 1 知

$$\lambda_k \ge x^H A x.$$

又因 $x \in V$，则有

$$x^H A x \ge \min_{v\in V, \|v\|=1} v^H A v.$$

进而有

$$\lambda_k \ge \min_{v\in V, \|v\|=1} v^H A v.$$

上式除了 $\dim(V) = k$ 外，没有任何限制，于是

$$\lambda_k \ge \max_{V, \dim(V)=k} \quad \min_{x\in V, \|x\|=1} x^H A x.$$

下面证明上面不等式的方向. 记 $W = \mathrm{span}\{u_1, \cdots, u_k\}$，由引理 1 知，若 $w \in W$，且 $\|w\| = 1$，有

$$w^H A w \ge \lambda_k = u_k^H A u_k.$$

因此，对 $\forall w \in W$，$\|w\| = 1$，有

$$\min_{w \in W, \|w\|=1} w^H A w \geq \lambda_k = u_k^H A u_k.$$

设 V 是 $\mathbf{C}^n$ 上任意的子空间，且 $\dim(V) = k$，则有

$$\max_{V, \dim(V)=k} \quad \min_{x \in V, \|x\|=1} x^H A x \geq \min_{w \in W, \|w\|=1} w^H A w \geq \lambda_k = u_k^H A u_k.$$

综上，不等式大于等于小于等于同时成立，即

$$\max_{V, \dim(V)=k} \quad \min_{x \in V, \|x\|=1} R(x) = \lambda_k.$$

同理，可证另外一个等式.

定理 3. ***(Weyl)*** *设* $A, B \in \mathbf{C}^{n \times n}$ *为* ***Hermitian*** *矩阵，则* $\forall k = 1, 2, \cdots, n$，*有*

$$\lambda_k(A) + \lambda_n(B) \leq \lambda_k(A+B) \leq \lambda_k(A) + \lambda_1(B).$$

证明： 对 $\forall x \in \mathbf{C}^n$，且 $\|x\| = 1$ 有

$$\lambda_n(B) \leq x^H B x \leq \lambda_1(B),$$

所以

$$\begin{aligned}
\lambda_k(A+B) &= \max_{V, \dim(V)=k} \quad \min_{x \in V, \|x\|=1} x^H (A+B) x \\
&= \max_{V, \dim(V)=k} \quad \min_{x \in V, \|x\|=1} \left(x^H A x + x^H B x\right) \\
&\geq \max_{V, \dim(V)=k} \quad \min_{x \in V, \|x\|=1} \left(x^H A x + \lambda_n(B)\right) \\
&= \lambda_k(A) + \lambda_n(B).
\end{aligned}$$

类似可证不等式另外一侧.

§1 矩阵序列与矩阵级数

程光辉

2020 年 5 月 10 日

设 $m \times n$ 的矩阵序列为 $\{A^{(k)}\}$，其中

$$A^{(k)} = \begin{bmatrix} a_{11}^{(k)} & a_{12}^{(k)} & \cdots & a_{1n}^{(k)} \\ a_{21}^{(k)} & a_{22}^{(k)} & \cdots & a_{2n}^{(k)} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}^{(k)} & a_{m2}^{(k)} & \cdots & a_{mn}^{(k)} \end{bmatrix}, \quad k = 1, 2, \cdots$$

定义 1 设 $A^{(k)} = (a_{ij}^{(k)}) \in \mathbf{C}^{m\times n}$, 若对 $1 \le i \le m, 1 \le j \le n$ 有 $\lim\limits_{k\to\infty} a_{ij}^{(k)} = a_{ij}$，则称矩阵序列 $\{A^{(k)}\}$ 收敛于 A，记为 $\lim\limits_{k\to\infty} A^{(k)} = A$.

定理 1 设 $\lim\limits_{k\to\infty} A^{(k)} = A$，$\lim\limits_{k\to\infty} B^{(k)} = B$, $\alpha, \beta \in \mathbf{C}$, 则

(1) $\lim\limits_{k\to\infty} (\alpha A^{(k)} + \beta B^{(k)}) = \alpha A + \beta B$,

(2) $\lim\limits_{k\to\infty} A^{(k)} B^{(k)} = AB$,

(3) 当 $A^{(k)}$ 与 A 都可逆时，$\lim\limits_{k\to\infty} (A^{(k)})^{-1} = A^{-1}$.

定理 2 设 $\|\cdot\|$ 是 $\mathbf{C}^{m\times n}$ 上的任一矩阵范数，$\mathbf{C}^{m\times n}$ 中矩阵序列 $\{A^{(k)}\}$ 收敛于 A 的充要条件是

$$\lim_{k\to\infty} \|A^{(k)} - A\| = 0.$$

定义 2 设 $A \in \mathbf{C}^{n\times n}$，若 $\lim\limits_{k\to\infty} A^k = O$，则称 A 为收敛矩阵.

定理 3 设 $A \in \mathbf{C}^{n\times n}$，则 A 为收敛矩阵的充要条件是 $r(A) < 1$.

证明： (必要性) 存在可逆矩阵 P，使得 A 相似于 Jordan 标准型，即

$$A = PJP^{-1} = P\text{diag}(J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_s}(\lambda_s))P^{-1},$$

则有

$$A^k = PJ^kP^{-1} = P\text{diag}(J_{r_1}^k(\lambda_1), J_{r_2}^k(\lambda_2), \cdots, J_{r_s}^k(\lambda_s))P^{-1}.$$

若 A 为收敛矩阵，则 $\lim\limits_{k\to\infty} A^k = O$，等价于 $\lim\limits_{k\to\infty} J_{r_i}^k(\lambda_i) = O, i = 1, 2, \cdots, s$. 即有 $J_{r_i}^k(\lambda_i)$ 的对角线元素 $|\lambda_i| < 1, i = 1, 2, \cdots, s$，则 $r(A) < 1$.

(充分性) $\lim\limits_{k\to\infty} J^k = O$ 等价于 $\lim\limits_{k\to\infty} J_{r_i}^k(\lambda_i) = O, i = 1, 2, \cdots, s$，而

$$J_{r_i}^k(\lambda_i) = \begin{bmatrix} f_k(\lambda_i) & f_k'(\lambda_i) & \cdots & \frac{f_k^{(r_i-1)}(\lambda_i)}{(r_i-1)!} \\ 0 & f_k(\lambda_i) & \cdots & \frac{f_k^{(r_i-2)}(\lambda_i)}{(r_i-2)!} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f_k(\lambda_i) \end{bmatrix}, \quad k > r_i,$$

其中 $f_k(\lambda_i) = \lambda_i^k$.

因为 r_i 为固定常数，当 $|\lambda_i| \le r(A) < 1$ 时，有

$$\lim_{k\to\infty} f_k^{(l)}(\lambda_i) = 0, \quad l = 0, 1, \cdots, r_i - 1.$$

定义 3 设矩阵序列 $\{A^{(k)}\}$，称

$$\sum_{k=1}^{\infty} A^{(k)} = A^{(1)} + A^{(2)} + \cdots + A^{(k)} + \cdots,$$

为矩阵级数，$S^{(N)} = \sum\limits_{k=1}^{N} A^{(k)}$ 为矩阵级数的部分和. 如果 $\lim\limits_{N\to\infty} S^{(N)} = S$，则称级数收敛. 不收敛的矩阵级数称为是发散的.

定义 4 设矩阵序列 $A^{(k)} = (a_{ij}^{(k)}) \in \mathbf{C}^{m\times n}$，如果 mn 个数项级数

$$\sum_{k=1}^{\infty} a_{ij}^{(k)}, i = 1, \cdots, m; j = 1, \cdots, n$$

都绝对收敛，则称矩阵级数 $\sum\limits_{k=1}^{\infty} A^{(k)}$ 绝对收敛.

定理 4 在 $\mathbf{C}^{m\times n}$ 中，矩阵级数 $\sum\limits_{k=1}^{\infty} A^{(k)}$ 绝对收敛的充要条件是正项级数 $\sum\limits_{k=1}^{\infty} \|A^{(k)}\|$ 收敛，其中 $\|\cdot\|$ 是任意矩阵范数.

证明： (必要性) 因为矩阵级数 $\sum\limits_{k=1}^{\infty} A^{(k)}$ 绝对收敛，则存在正数 M，使得

$$\sum_{k=1}^{N} |a_{ij}^{(k)}| \le M,\ i = 1, 2, \cdots, m;\ j = 1, 2, \cdots, n.$$

从而

$$\sum_{k=1}^{N} \|A^{(k)}\|_{m_1} = \sum_{k=1}^{N} \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}^{(k)}| \right) \le mnM,$$

故 $\sum\limits_{k=1}^{N}\|\boldsymbol{A}^{(k)}\|_{m_1}$ 收敛. 再根据范数的等价性和正项级数的比较判别法知,正项级数 $\sum\limits_{k=1}^{\infty}\|\boldsymbol{A}^{(k)}\|$ 收敛.

(充分性) 因正项级数 $\sum\limits_{k=1}^{\infty}\|\boldsymbol{A}^{(k)}\|$ 收敛，故 $\sum\limits_{k=1}^{\infty}\|\boldsymbol{A}^{(k)}\|_{m_1}$ 收敛，即有

$$|a_{ij}^{(k)}| \leq \|\boldsymbol{A}^{(k)}\|_{m_1},\ i=1,2,\cdots,m;\ j=1,2,\cdots,n,$$

故由级数比较判别法知，$\sum\limits_{k=1}^{\infty}a_{ij}^{(k)}, i=1,\cdots,m; j=1,\cdots,n$ 绝对收敛，得证.

定理 5 n 方阵 $\boldsymbol{A}$ 的幂级数 *(Neumann 级数)*

$$\sum_{k=0}^{\infty}\boldsymbol{A}^k=\boldsymbol{E}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots$$

收敛的充要条件是 $r(\boldsymbol{A})<1$，且收敛时，其和为 $(\boldsymbol{E}-\boldsymbol{A})^{-1}$.

证明： (必要性) 因为 $\sum\limits_{k=0}^{\infty}\boldsymbol{A}^k$ 收敛，则

$$\sum_{k=0}^{\infty}(\boldsymbol{A}^k)_{ij},\ i,j=1,2,\cdots,n,$$

收敛，进而 $\lim\limits_{k\to\infty}(\boldsymbol{A}^k)_{ij}=0,\ i,j=1,2,\cdots,n$，即 $\lim\limits_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{O}$，于是 $r(\boldsymbol{A})<1$.

(充分性) 因为 $r(\boldsymbol{A})<1$，故 $\boldsymbol{E}-\boldsymbol{A}$ 可逆. 又因

$$(\boldsymbol{E}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k)(\boldsymbol{E}-\boldsymbol{A})=\boldsymbol{E}-\boldsymbol{A}^{k+1},$$

所以

$$\boldsymbol{E}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k=(\boldsymbol{E}-\boldsymbol{A})^{-1}-\boldsymbol{A}^{k+1}(\boldsymbol{E}-\boldsymbol{A})^{-1}.$$

当 $k\to\infty$ 时，有 $\boldsymbol{A}^{k+1}(\boldsymbol{E}-\boldsymbol{A})^{-1}\to\boldsymbol{O}$，即

$$\lim_{n\to\infty}\sum_{k=0}^{n}\boldsymbol{A}^k=(\boldsymbol{E}-\boldsymbol{A})^{-1}.$$

定理 6 设幂级数

$$f(z)=\sum_{k=0}^{\infty}a_kz^k$$

的收敛半径为 r，如果方阵 $\boldsymbol{A}$ 满足 $r(\boldsymbol{A})<r$，则矩阵幂级数 $\sum\limits_{k=0}^{\infty}c_k\boldsymbol{A}^k$ 绝对收敛；若 $r(\boldsymbol{A})>r$ 则发散.

§2 矩阵函数

程光辉

2020 年 5 月 10 日

定义 1 设幂级数 $\sum\limits_{k=0}^{\infty} a_k z^k$ 的收敛半径为 r，且当 $|z| < r$ 时，幂级数收敛于 $f(z)$，即

$$f(z) = \sum_{k=0}^{\infty} a_k z^k, \quad |z| < r.$$

如果方阵 $A \in \mathbf{C}^{n\times n}$ 满足 $r(A) < r$，则称收敛的矩阵幂级数 $\sum\limits_{k=0}^{\infty} a_k A^k$ 的和为矩阵函数，记为 $f(A)$，即

$$f(A) = \sum_{k=0}^{\infty} a_k A^k.$$

把 $f(A)$ 的方阵 A 换为 At，t 为参数，则有

$$f(At) = \sum_{k=0}^{\infty} a_k (At)^k.$$

若 $A \in \mathbf{C}^{n\times n}$，则常用的矩阵函数：

(1) $e^A = \sum\limits_{k=0}^{\infty} \frac{1}{k!} A^k$,

(2) $\sin A = \sum\limits_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} A^{2k+1}$,

(3) $\cos A = \sum\limits_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} A^{2k}$,

(4) $(E - A)^{-1} = \sum\limits_{k=0}^{\infty} A^k, \quad r(A) < 1$,

(5) $\ln(E + A) = \sum\limits_{k=0}^{\infty} \frac{(-1)^k}{(k+1)} A^{k+1}, \quad r(A) < 1$.

常用矩阵函数的性质：

(1) $e^{iA} = \cos A + i \sin A$,

(2) $\cos A = \frac{1}{2}(e^{iA} + e^{-iA})$,

(3) $\sin A = \frac{1}{2i}(e^{iA} - e^{-iA})$,

(4) $\cos(-A) = \cos A, \quad \sin(-A) = -\sin A$.

矩阵函数的计算方法：

(1) 利用相似对角化 (只对单纯矩阵有用)

(2) Jordan 标准型法 (计算困难)

(3) 数项级数求和法 (不具有通用性)

1. 利用相似对角化

设 $A = P\mathrm{diag}\left(\lambda_1, \lambda_2, \cdots, \lambda_n\right) P^{-1} = PDP^{-1}$，则

$$\begin{aligned}
f(A) &= \sum_{k=0}^{\infty} a_k A^k \\
&= P\left(\sum_{k=0}^{\infty} a_k D^k\right) P^{-1} \\
&= P\mathrm{diag}\left(\sum_{k=0}^{\infty} a_k \lambda_1^k, \cdots, \sum_{k=0}^{\infty} a_k \lambda_n^k\right) P^{-1} \\
&= P\mathrm{diag}\left(f(\lambda_1), \cdots, f(\lambda_n)\right) P^{-1}.
\end{aligned}$$

例 1 设

$$A = \begin{bmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{bmatrix},$$

求 e^{At}.

解：计算步骤为

- 计算特征值： $\lambda_1 = -2,\ \lambda_2 = \lambda_3 = 1$.
- 计算特征向量： $x_1 = [-1\ 1\ 1]^T, \quad x_2 = [-2\ 1\ 0]^T, \quad x_3 = [0\ 0\ 1]^T$.
- 构造矩阵：

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

- 写出计算值：

$$
\begin{aligned}
e^{At} &= P\begin{bmatrix} e^{-2t} & 0 & 0 \\ 0 & e^t & 0 \\ 0 & 0 & e^t \end{bmatrix} P^{-1} \\
&= \begin{bmatrix} 2e^t - e^{-2t} & 2e^t - 2e^{-2t} & 0 \\ e^{2t} - e^{-t} & 2e^{-2t} - e^t & 0 \\ e^{-2t} - e^t & 2e^{-2t} - 2e^t & e^t \end{bmatrix}.
\end{aligned}
$$

2. Jordan 标准型方法

设 $P^{-1}AP = J = \mathrm{diag}(J_1, \cdots, J_s)$，则 $A = PJP^{-1} = P\mathrm{diag}(J_1, \cdots, J_s)P^{-1}$，进而

$$
f(A) = P\left(\sum_{k=0}^{\infty} a_k J^k\right) P^{-1} = P\mathrm{diag}\left(f(J_1), \cdots, f(J_s)\right) P^{-1},
$$

其中

$$
\begin{aligned}
f(J_i) &= \sum_{k=0}^{\infty} a_k J_i^k \\
&= \sum_{k=0}^{\infty} a_k \begin{bmatrix} \lambda_i^k & C_k^1 \lambda_i^{k-1} & \cdots & C_k^{r_i-1}\lambda_i^{k-(r_i-1)} \\ & \lambda_i^k & & \vdots \\ & & \ddots & C_k^1 \lambda_i^{k-1} \\ & & & \lambda_i^k \end{bmatrix} \\
&= \begin{bmatrix} f(\lambda_i) & \frac{1}{1!}f'(\lambda_i) & \cdots & \frac{1}{(r_i-1)!}f^{(r_i-1)}(\lambda_i) \\ & f(\lambda_i) & \cdots & \frac{1}{(r_i-2)!}f^{(r_i-2)}(\lambda_i) \\ & & \ddots & \vdots \\ & & & f(\lambda_i) \end{bmatrix},
\end{aligned}
$$

其中 r_i 是特征值 λ_i 的代数重数.

例 2 设

$$
A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},
$$

求 $\sin A$.

解： (1) 化为 Jordan 标准型：$J_1 = 1$, $\quad J_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

(2) 计算 $\sin J_i$：

$$\sin J_1=\sin 1,\quad \sin J_2=\begin{bmatrix}\sin 1 & \frac{1}{1!}\cos 1\\ 0 & \sin 1\end{bmatrix},$$

则

$$\sin A=\begin{bmatrix}\sin 1 & 0 & 0\\ 0 & \sin 1 & \cos 1\\ 0 & 0 & \sin 1\end{bmatrix}.$$

例 3 求 $\sin A$，其中

$$A=\begin{bmatrix}\pi & 0 & 0 & 0\\ 0 & -\pi & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\end{bmatrix}.$$

解： (1) 化为 Jordan 标准型：$J_1=\pi,\quad J_2=-\pi,\quad J_3=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$.

(2) 计算 $\sin J_i$：

$$\sin J_1=0,\quad \sin J_2=0,\quad \sin J_3=\begin{bmatrix}\sin 0 & \frac{1}{1!}\cos 0\\ 0 & \sin 0\end{bmatrix}=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix},$$

则

$$\sin A=\begin{bmatrix}0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\end{bmatrix}.$$

3. 数项级数求和法

定理 1 哈密尔顿-凯莱定理： 设 A 是数域 P 上的一个 $n\times n$ 矩阵，$f(\lambda)=|\lambda E-A|=\lambda^n-b_{n-1}\lambda^{n-1}-\cdots-b_1\lambda-b_0$ 是 A 的特征多项式，则

$$f(A)=A^n-b_{n-1}A^{n-1}-\cdots-b_1A-b_0E=O.$$

于是有

$$A^n=b_{n-1}A^{n-1}+\cdots+b_1A+b_0E,$$

进而

$$\begin{cases}A^{n+1}=b_{n-1}^{(1)}A^{n-1}+\cdots+b_1^{(1)}A+b_0^{(1)}E\\ \cdots\cdots\\ A^{n+l}=b_{n-1}^{(l)}A^{n-1}+\cdots+b_1^{(l)}A+b_0^{(l)}E\\ \cdots\cdots\end{cases}$$

所以

$$
\begin{aligned}
f(A)=&\sum_{k=0}^{\infty}c_kA^k\\
=&(c_0E+c_1A+\cdots+c_{n-1}A^{n-1})\\
&+c_n(b_{n-1}A^{n-1}+\cdots+b_1A+b_0E)+\cdots\\
=&(c_0+\sum_{l=0}^{\infty}c_{n+l}b_0^{(l)})E+(c_1+\sum_{l=0}^{\infty}c_{n+l}b_1^{(l)})A\\
&+(c_{n-1}+\sum_{l=0}^{\infty}c_{n+l}b_{n-1}^{(l)})A^{n-1}.
\end{aligned}
$$

例 4 求 $\sin A$，其中

$$
A=\begin{bmatrix}\pi & 0 & 0 & 0\\ 0 & -\pi & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\end{bmatrix}.
$$

解：特征多项式为 $\det(\lambda E-A)=\lambda^4-\pi^2\lambda^2$，所以有 $A^4=\pi^2A^2$，$A^5=\pi^2A^3$，$A^7=\pi^4A^3$，$\cdots$，于是有

$$
\begin{aligned}
\sin A&=A-\frac{1}{3!}A^3+\frac{1}{5!}A^5-\frac{1}{7!}A^7+\frac{1}{9!}A^9-\cdots\\
&=A-\frac{1}{3!}A^3+\frac{1}{5!}\pi^2A^3-\frac{1}{7!}\pi^4A^3+\frac{1}{9!}\pi^6A^3-\cdots\\
&=A-\left(\frac{1}{3!}-\frac{1}{5!}\pi^2+\frac{1}{7!}\pi^4-\frac{1}{9!}\pi^6+\cdots\right)A^3\\
&=A+\frac{\sin\pi-\pi}{\pi^3}A^3\\
&=\begin{bmatrix}0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\end{bmatrix}.
\end{aligned}
$$

矩阵函数的一些性质：

(1) 如果 $AB=BA$，则 $e^Ae^B=e^Be^A=e^{A+B}$.

(2) $e^Ae^{-A}=E$.

(3) $(e^A)^m=e^{mA}$.

(4) 如果 $\boldsymbol{AB}=\boldsymbol{BA}$，则

$$\begin{aligned}\cos(\boldsymbol{A}+\boldsymbol{B})&=\cos\boldsymbol{A}\cos\boldsymbol{B}-\sin\boldsymbol{A}\sin\boldsymbol{B},\\ \sin(\boldsymbol{A}+\boldsymbol{B})&=\sin\boldsymbol{A}\cos\boldsymbol{B}+\cos\boldsymbol{A}\sin\boldsymbol{B}.\end{aligned}$$

§1 单边逆矩阵

程光辉

2019 年 12 月 2 日

定义 1 设 $\boldsymbol{A} \in \mathbf{C}^{m \times n}$，如果有 $\boldsymbol{G} \in \mathbf{C}^{n \times m}$，使得

$$\boldsymbol{GA} = \boldsymbol{E}_n,$$

则称 $\boldsymbol{G}$ 为 $\boldsymbol{A}$ 的左逆矩阵，记为 $\boldsymbol{G} = \boldsymbol{A}_L^{-1}$.

如果有 $\boldsymbol{G} \in \mathbf{C}^{n \times m}$，使得

$$\boldsymbol{AG} = \boldsymbol{E}_m,$$

则称 $\boldsymbol{G}$ 为 $\boldsymbol{A}$ 的右逆矩阵，记为 $\boldsymbol{G} = \boldsymbol{A}_R^{-1}$.

定理 1 设 $\boldsymbol{A} \in \mathbf{C}^{m \times n}$，则

(1) $\boldsymbol{A}$ *左可逆的充要条件是* $\boldsymbol{A}$ *为列满秩矩阵;*

(2) $\boldsymbol{A}$ *右可逆的充要条件是* $\boldsymbol{A}$ *为行满秩矩阵.*

证明： (1) **充分性：** 因 $\boldsymbol{A}$ 为列满秩矩阵，则 $\boldsymbol{A}^H\boldsymbol{A}$ 为满秩矩阵，进而

$$\left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}\boldsymbol{A}^H\boldsymbol{A} = \boldsymbol{GA} = \boldsymbol{E}_n,$$

其中 $\boldsymbol{G} = \left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}\boldsymbol{A}^H$ 为矩阵 $\boldsymbol{A}$ 的左逆.

必要性： 因为 $\boldsymbol{A}_L^{-1}\boldsymbol{A} = \boldsymbol{E}_n$，则

$$\mathbf{rank}(\boldsymbol{A}) \geq \mathbf{rank}(\boldsymbol{A}_L^{-1}\boldsymbol{A}) = \mathbf{rank}(\boldsymbol{E}_n) = \boldsymbol{n},$$

因此，$\mathbf{rank}(\boldsymbol{A}) = \boldsymbol{n}$，即 $\boldsymbol{A}$ 为列满秩矩阵.

推论 1 设 $\boldsymbol{A} \in \mathbf{C}^{m \times n}$，则

(1) $\boldsymbol{A}$ *左可逆的充要条件是* $\mathbf{N}(\boldsymbol{A}) = \{\mathbf{0}\}$*;*

(2) $\boldsymbol{A}$ *右可逆的充要条件是* $\mathbf{R}(\boldsymbol{A}) = \mathbf{C}^m$*.*

证明： (1) **充分性：**因为 $\mathbf{N}(\boldsymbol{A})=\{\mathbf{0}\}$，则 $\boldsymbol{Ax}=\mathbf{0}$ 只有零解，系数矩阵列满秩，即 $\boldsymbol{A}$ 左可逆的.

必要性： $\boldsymbol{A}$ 左可逆，则 $\boldsymbol{A}_L^{-1}\boldsymbol{A}=\boldsymbol{E}$，对 $\forall \boldsymbol{x}\in \mathbf{N}(\boldsymbol{A})$，有

$$\boldsymbol{x}=\boldsymbol{Ex}=\boldsymbol{A}_L^{-1}\boldsymbol{Ax}=\boldsymbol{A}_L^{-1}\mathbf{0}=\mathbf{0},$$

即 $\mathbf{N}(\boldsymbol{A})=\{\mathbf{0}\}$.

初等变换求左 (右) 逆矩阵：

(1) $\boldsymbol{P}\begin{bmatrix}\boldsymbol{A} & \boldsymbol{E}_m\end{bmatrix}=\begin{bmatrix}\boldsymbol{E}_n & \boldsymbol{G}\\ \boldsymbol{O} & \star\end{bmatrix}$;

(2) $\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{E}_n\end{bmatrix}\boldsymbol{Q}=\begin{bmatrix}\boldsymbol{E}_m & \boldsymbol{O}\\ \boldsymbol{G} & \star\end{bmatrix}$.

例 1 设矩阵 $\boldsymbol{A}$ 为

$$\boldsymbol{A}=\begin{bmatrix}1 & -2\\ 0 & 1\\ 0 & 0\end{bmatrix},$$

求 $\boldsymbol{A}$ 的一个左逆矩阵 $\boldsymbol{A}_L^{-1}$.

解：经行初等变换，有

$$\begin{bmatrix}\boldsymbol{A} & \boldsymbol{E}_3\end{bmatrix}=\begin{bmatrix}1 & -2 & 1 & 0 & 0\\ 0 & 1 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 1\end{bmatrix}\to\begin{bmatrix}1 & 0 & 1 & 2 & 0\\ 0 & 1 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 1\end{bmatrix},$$

因此，$\boldsymbol{A}_L^{-1}=\begin{bmatrix}1 & 2 & 0\\ 0 & 1 & 0\end{bmatrix}$.

例 2 设矩阵 $\boldsymbol{A}$ 为

$$\boldsymbol{A}=\begin{bmatrix}1 & 2 & -1\\ 0 & -1 & 2\end{bmatrix},$$

求 $\boldsymbol{A}$ 的一个右逆矩阵 $\boldsymbol{A}_R^{-1}$.

解：经列初等变换，有

$$\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{E}_3\end{bmatrix}=\begin{bmatrix}1 & 2 & -1\\ 0 & -1 & 2\\ 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\to\begin{bmatrix}1 & 0 & 0\\ 0 & -1 & 2\\ 1 & -2 & 1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\to\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 1 & 2 & -3\\ 0 & -1 & 2\\ 0 & 0 & 1\end{bmatrix},$$

因此，$A_R^{-1} = \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ 0 & 0 \end{bmatrix}$.

定理 2 设 $A \in \mathbf{C}^{m\times n}$ 是左可逆矩阵，则

$$G = \begin{bmatrix} A_1^{-1} - BA_2A_1^{-1} & B \end{bmatrix} P$$

是 A 的左逆矩阵，其中 $B \in \mathbf{C}^{n\times(m-n)}$ 的任意矩阵，行初等变换矩阵 P 满足 $PA = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$，$A_1$ 是 n 阶可逆矩阵.

证明： 直接验证，即

$$GA = \begin{bmatrix} A_1^{-1} - BA_2A_1^{-1} & B \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = E_n.$$

定理 3 设 $A \in \mathbf{C}^{m\times n}$ 是右可逆矩阵，则

$$G = Q \begin{bmatrix} A_1^{-1} - A_1^{-1}A_2D \\ D \end{bmatrix}$$

是 A 的右逆矩阵，其中 $D \in \mathbf{C}^{(n-m)\times m}$ 的任意矩阵，列初等变换矩阵 Q 满足 $AQ = \begin{bmatrix} A_1 & A_2 \end{bmatrix}$，$A_1$ 是 m 阶可逆矩阵.

定理 4 设 $A \in \mathbf{C}^{m\times n}$ 是左可逆矩阵，A_L^{-1} 是 A 的左逆矩阵，则方程组 $Ax = b$ 有解的充要条件是

$$(E_m - AA_L^{-1})b = 0. \tag{1}$$

若 (1) 成立，则方程组 $Ax = b$ 有唯一解

$$x = \left(A^H A\right)^{-1} A^H b.$$

证明： (必要性) 设 x_0 是方程组 $Ax = b$ 的解，则

$$(AA_L^{-1})b = (AA_L^{-1})(Ax_0) = A(A_L^{-1}A)x_0 = AE_nx_0 = Ax_0 = b,$$

进而有 $(E_m - AA_L^{-1})b = 0$.

(充分性) 因为 $(E_m - AA_L^{-1})b = 0$，故有 $AA_L^{-1}b = b$，即方程组 $Ax = b$ 有解 $A_L^{-1}b$.

(唯一性) 设 x_0，x_1 是 $Ax = b$ 的解，则 $A(x_0 - x_1) = b - b = 0$. 又因为 A 为列满秩矩阵，故只有零解，即 $x_0 = x_1$.

定理 5 设 $\boldsymbol{A} \in \mathbf{C}^{m \times n}$ 是右可逆矩阵，则方程组 $\boldsymbol{A x}=\boldsymbol{b}$ 对任何 $\boldsymbol{b} \in \mathbf{C}^{m}$ 都有解，若 $\boldsymbol{b} \neq \mathbf{0}$，则方程组的解可表示为

$$\boldsymbol{x}=\boldsymbol{A}_{R}^{-1} \boldsymbol{b}.$$

证明： 因为 $\boldsymbol{b}=\boldsymbol{E}_{m} \boldsymbol{b}=\boldsymbol{A} \boldsymbol{A}_{R}^{-1} \boldsymbol{b}=(\boldsymbol{A} \boldsymbol{A}_{R}^{-1}) \boldsymbol{b}$，故 $\boldsymbol{x}=\boldsymbol{A}_{R}^{-1} \boldsymbol{b}$ 是 $\boldsymbol{A x}=\boldsymbol{b}$ 的解.

§2 广义逆矩阵 A^-

程光辉

2019 年 12 月 6 日

定义 1 设 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，如果存在矩阵 $\boldsymbol{G} \in \mathbf{C}^{n\times m}$，使得

$$\boldsymbol{AGb} = \boldsymbol{b}, \quad \forall \boldsymbol{b} \in \mathbf{R}(\boldsymbol{A}),$$

则称 $\boldsymbol{G}$ 为 $\boldsymbol{A}$ 的广义逆矩阵，记为 $\boldsymbol{G} = \boldsymbol{A}^-$.

定理 1 设 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，则 $\boldsymbol{A}$ 存在广义逆矩阵的充要条件是存在 $\boldsymbol{G} \in \mathbf{C}^{n\times m}$，满足

$$\boldsymbol{AGA} = \boldsymbol{A}.$$

证明： **(必要性)** 对 $\forall \boldsymbol{u} \in \mathbf{C}^n$，则 $\boldsymbol{b} = \boldsymbol{Au} \in \mathbf{R}(\boldsymbol{A})$. 因 $\boldsymbol{AGb} = \boldsymbol{b}$，故

$$\boldsymbol{AGAu} = \boldsymbol{AGb} = \boldsymbol{b} = \boldsymbol{Au},$$

由于 $\boldsymbol{u}$ 的任意性 (可取单位矩阵的列向量)，故 $\boldsymbol{AGA} = \boldsymbol{A}$.

(充分性) 对 $\forall \boldsymbol{b} \in \mathbf{R}(\boldsymbol{A})$，则存在 $\boldsymbol{u} \in \mathbf{C}^n$ 使得 $\boldsymbol{b} = \boldsymbol{Au}$. 因

$$\boldsymbol{b} = \boldsymbol{Au} = \boldsymbol{AGAu} = \boldsymbol{AGb},$$

故 $\boldsymbol{G}$ 为 $\boldsymbol{A}$ 的广义逆矩阵.

推论 1 设 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，且 $\boldsymbol{A}^-$ 是 $\boldsymbol{A}$ 的一个广义逆矩阵，则

$$\mathbf{rank}(\boldsymbol{A}^-) \geq \mathbf{rank}(\boldsymbol{A}).$$

证明： 因为

$$\mathbf{rank}(\boldsymbol{A}) = \mathbf{rank}(\boldsymbol{AA^-A}) \leq \mathbf{rank}(\boldsymbol{AA^-}) \leq \mathbf{rank}(\boldsymbol{A}^-),$$

故得证.

定义 2 $\boldsymbol{A}\{1\} = \{\boldsymbol{G} | \boldsymbol{AGA} = \boldsymbol{A},\ \forall \boldsymbol{G} \in \mathbf{C}^{n\times m}\}$.

定理 2 设 $A \in \mathbf{C}^{m\times n}$，且 A^- 是 A 的任意广义逆矩阵，则

$$\begin{aligned} A\{1\} &= \{G | G = A^- + U - A^-AUAA^-, \ \forall U \in \mathbf{C}^{n\times m}\} \\ &= \{G | G = A^- + (E_n - A^-A)V + W(E_m - AA^-), \ \forall V, W \in \mathbf{C}^{n\times m}\}. \end{aligned}$$

证明：若两个集合互相包含，则这两个集合相等.

对 $\forall G \in A\{1\}$，则 $AGA = A$，于是有

$$G = A^- + G - A^- - A^-A(G - A^-)AA^- = A^- + U - A^-AUAA^-,$$

其中 $U = G - A^-$，进而 $A\{1\} \subset \{G | G = A^- + U - A^-AUAA^-, \ \forall U \in \mathbf{C}^{n\times m}\}$.

对 $\forall M \in \{G | G = A^- + U - A^-AUAA^-, \ \forall U \in \mathbf{C}^{n\times m}\}$，则存在 $U \in \mathbf{C}^{n\times m}$，使得

$$M = A^- + U - A^-AUAA^-,$$

进而

$$\begin{aligned} AMA &= A(A^- + U - A^-AUAA^-)A \\ &= AA^-A + AUA - AA^-AUAA^-A \\ &= A + AUA - AUA \\ &= A. \end{aligned}$$

因此，$\{G | G = A^- + U - A^-AUAA^-, \ \forall U \in \mathbf{C}^{n\times m}\} \subset A\{1\}$，得证.

定理 3 设 $A \in \mathbf{C}^{m\times n}$，$\lambda \in \mathbf{C}$ 非零，则

(1) $(A^T)^- = (A^-)^T$，$(A^H)^- = (A^-)^H$；

(2) AA^-，A^-A 都是幂等矩阵，且 $\operatorname{rank}(A) = \operatorname{rank}(AA^-) = \operatorname{rank}(A^-A)$;

(3) $\lambda^{-1}A^-$ 为 λA 的广义逆矩阵;

(4) 设 $S \in \mathbf{C}^{m\times m}$ 是可逆矩阵，$T \in \mathbf{C}^{n\times n}$ 是可逆矩阵，且 $B = SAT$，则 $T^{-1}A^-S^{-1}$ 是 B 的广义逆矩阵;

(5) $\mathrm{R}(AA^-) = \mathrm{R}(A)$，$\mathrm{N}(A^-A) = \mathrm{N}(A)$.

证明：(1) 因为 $AA^-A = A$，所以 $A^T = A^T(A^-)^TA^T$，即 $(A^T)^- = (A^-)^T$.

同理，$(A^H)^- = (A^-)^H$.

(2) 因 $(AA^-)^2 = AA^-AA^- = AA^-$，故是幂等矩阵.

因 $\operatorname{rank}(A) \geq \operatorname{rank}(AA^-) \geq \operatorname{rank}(AA^-A) = \operatorname{rank}(A)$，故

$$\operatorname{rank}(A) = \operatorname{rank}(AA^-).$$

同理，可证 $\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A}^-\boldsymbol{A})$.

(3) $(\lambda\boldsymbol{A})(\lambda^{-1}\boldsymbol{A}^-)(\lambda\boldsymbol{A}) = (\lambda\lambda^{-1}\lambda)\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A} = \lambda\boldsymbol{A}$.

(4) 因 $\boldsymbol{B}\boldsymbol{T}^{-1}\boldsymbol{A}^-\boldsymbol{S}^{-1}\boldsymbol{B} = \boldsymbol{S}\boldsymbol{A}\boldsymbol{T}\boldsymbol{T}^{-1}\boldsymbol{A}^-\boldsymbol{S}^{-1}\boldsymbol{S}\boldsymbol{A}\boldsymbol{T} = \boldsymbol{S}\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A}\boldsymbol{T} = \boldsymbol{S}\boldsymbol{A}\boldsymbol{T} = \boldsymbol{B}$，则 $\boldsymbol{T}^{-1}\boldsymbol{A}^-\boldsymbol{S}^{-1}$ 是 $\boldsymbol{B}$ 的广义逆矩阵.

(5) 显然有 $\mathrm{R}(\boldsymbol{A}\boldsymbol{A}^-) \subset \mathrm{R}(\boldsymbol{A})$，$\mathrm{N}(\boldsymbol{A}) \subset \mathrm{N}(\boldsymbol{A}^-\boldsymbol{A})$，又因

$$\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A}\boldsymbol{A}^-) = \mathrm{rank}(\boldsymbol{A}^-\boldsymbol{A}),$$

故 $\mathrm{R}(\boldsymbol{A}\boldsymbol{A}^-)$ 和 $\mathrm{R}(\boldsymbol{A})$ 的基相同，$\mathrm{N}(\boldsymbol{A})$ 和 $\mathrm{N}(\boldsymbol{A}^-\boldsymbol{A})$ 的基相同，即 $\mathrm{R}(\boldsymbol{A}\boldsymbol{A}^-) = \mathrm{R}(\boldsymbol{A})$，$\mathrm{N}(\boldsymbol{A}^-\boldsymbol{A}) = \mathrm{N}(\boldsymbol{A})$.

(另证：) 对 $\forall \boldsymbol{b} \in \mathrm{R}(\boldsymbol{A})$，则存在 $\boldsymbol{x} \in \mathrm{C}^n$，使得 $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$. 因为 $\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A} = \boldsymbol{A}$，则有

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{A}^-\boldsymbol{y} \in \mathrm{R}(\boldsymbol{A}\boldsymbol{A}^-),$$

其中 $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$. 由 $\boldsymbol{b}$ 的任意性，有 $\mathrm{R}(\boldsymbol{A}\boldsymbol{A}^-) \supset \mathrm{R}(\boldsymbol{A})$.

推论 2 设 $\boldsymbol{A} \in \mathrm{C}^{m\times n}$，则

(1) $\mathrm{rank}(\boldsymbol{A}) = n$ 的充要条件是 $\boldsymbol{A}^-\boldsymbol{A} = \boldsymbol{E}_n$;

(2) $\mathrm{rank}(\boldsymbol{A}) = m$ 的充要条件是 $\boldsymbol{A}\boldsymbol{A}^- = \boldsymbol{E}_m$.

证明： (1) **(充分性)** 由定理 3(2) 知，$\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A}^-\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{E}_n) = n$.

(必要性) 因为 $\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A}^-\boldsymbol{A}) = n$，则 $\boldsymbol{A}^-\boldsymbol{A}$ 是 n 阶可逆矩阵，即有

$$\boldsymbol{E}_n = (\boldsymbol{A}^-\boldsymbol{A})(\boldsymbol{A}^-\boldsymbol{A})^{-1} = \boldsymbol{A}^-(\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A})(\boldsymbol{A}^-\boldsymbol{A})^{-1} = (\boldsymbol{A}^-\boldsymbol{A})(\boldsymbol{A}^-\boldsymbol{A})(\boldsymbol{A}^-\boldsymbol{A})^{-1} = \boldsymbol{A}^-\boldsymbol{A}.$$

引理 1 设 $\boldsymbol{A} \in \mathrm{C}^{m\times n}$，且 $\boldsymbol{P} \in \mathrm{C}^{m\times m}$ 和 $\boldsymbol{Q} \in \mathrm{C}^{n\times n}$ 都是可逆矩阵，则

$$\boldsymbol{Q}(\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q})^-\boldsymbol{P} \in \boldsymbol{A}\{1\}.$$

证明: 因 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$ 都是可逆矩阵，$\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q}(\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q})^-\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{P}\boldsymbol{A}\boldsymbol{Q}$，所以 $\boldsymbol{A}\boldsymbol{Q}(\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q})^-\boldsymbol{P}\boldsymbol{A} = \boldsymbol{A}$，即

$$\boldsymbol{Q}(\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q})^-\boldsymbol{P} \in \boldsymbol{A}\{1\}.$$

引理 2 设 $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_{22} \end{bmatrix} \in \mathrm{C}^{m\times n}$，则存在 $\boldsymbol{X}_{12}, \boldsymbol{X}_{21}$ 满足 $\boldsymbol{A}_{11}\boldsymbol{X}_{12}\boldsymbol{A}_{22} = \boldsymbol{O}$，$\boldsymbol{A}_{22}\boldsymbol{X}_{21}\boldsymbol{A}_{11} = \boldsymbol{O}$，使得

$$\begin{bmatrix} \boldsymbol{A}_{11}^- & \boldsymbol{X}_{12} \\ \boldsymbol{X}_{21} & \boldsymbol{A}_{22}^- \end{bmatrix} \in \boldsymbol{A}\{1\}.$$

证明：直接验证即可.

$$\begin{aligned}
\boldsymbol{A}\begin{bmatrix}\boldsymbol{A}_{11}^{-} & \boldsymbol{X}_{12}\\ \boldsymbol{X}_{21} & \boldsymbol{A}_{22}^{-}\end{bmatrix}\boldsymbol{A} &= \begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{A}_{22}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}_{11}^{-} & \boldsymbol{X}_{12}\\ \boldsymbol{X}_{21} & \boldsymbol{A}_{22}^{-}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{A}_{22}\end{bmatrix}\\
&= \begin{bmatrix}\boldsymbol{A}_{11}\boldsymbol{A}_{11}^{-} & \boldsymbol{A}_{11}\boldsymbol{X}_{12}\\ \boldsymbol{A}_{22}\boldsymbol{X}_{21} & \boldsymbol{A}_{22}\boldsymbol{A}_{22}^{-}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{A}_{22}\end{bmatrix}\\
&= \begin{bmatrix}\boldsymbol{A}_{11}\boldsymbol{A}_{11}^{-}\boldsymbol{A}_{11} & \boldsymbol{A}_{11}\boldsymbol{X}_{12}\boldsymbol{A}_{22}\\ \boldsymbol{A}_{22}\boldsymbol{X}_{21}\boldsymbol{A}_{11} & \boldsymbol{A}_{22}\boldsymbol{A}_{22}^{-}\boldsymbol{A}_{22}\end{bmatrix}\\
&= \begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{A}_{22}\end{bmatrix}\\
&= \boldsymbol{A}.
\end{aligned}$$

定理 4 设 $\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix}\in\mathbf{C}^{m\times n}$，有

(1) 如果 $\boldsymbol{A}_{11}^{-1}$ 存在，则存在 $\boldsymbol{X}_{12}$，$\boldsymbol{X}_{21}$ 满足

$$\begin{aligned}
&\boldsymbol{X}_{12}(\boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})=\boldsymbol{O},\\
&(\boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})\boldsymbol{X}_{21}=\boldsymbol{O},
\end{aligned}$$

使得

$$\begin{bmatrix}\boldsymbol{E}_r & -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\\ \boldsymbol{O} & \boldsymbol{E}_{n-r}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}_{11}^{-1} & \boldsymbol{X}_{12}\\ \boldsymbol{X}_{21} & (\boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})^{-}\end{bmatrix}\begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{O}\\ -\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & \boldsymbol{E}_{m-r}\end{bmatrix}\in\boldsymbol{A}\{1\}.$$

(2) 如果 $\boldsymbol{A}_{22}^{-1}$ 存在，则存在 $\boldsymbol{Y}_{12}$，$\boldsymbol{Y}_{21}$ 满足

$$\begin{aligned}
&\boldsymbol{Y}_{21}(\boldsymbol{A}_{11}-\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21})=\boldsymbol{O},\\
&(\boldsymbol{A}_{11}-\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21})\boldsymbol{Y}_{12}=\boldsymbol{O},
\end{aligned}$$

使得

$$\begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{O}\\ -\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21} & \boldsymbol{E}_{n-r}\end{bmatrix}\begin{bmatrix}(\boldsymbol{A}_{11}-\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21})^{-} & \boldsymbol{Y}_{12}\\ \boldsymbol{Y}_{21} & \boldsymbol{A}_{22}^{-1}\end{bmatrix}\begin{bmatrix}\boldsymbol{E}_r & -\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\\ \boldsymbol{O} & \boldsymbol{E}_{m-r}\end{bmatrix}\in\boldsymbol{A}\{1\}.$$

证明：(1) 因为

$$\begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{O}\\ -\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & \boldsymbol{E}_{m-r}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix}\begin{bmatrix}\boldsymbol{E}_r & -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\\ \boldsymbol{O} & \boldsymbol{E}_{n-r}\end{bmatrix}=\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\end{bmatrix},$$

利用引理 1 和引理 2，即可得证.

(2) 同 (1) 的证明.

§3 自反广义逆矩阵

程光辉

2020 年 5 月 15 日

定义 1 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，如果存在矩阵 $\boldsymbol{G}\in\mathbf{C}^{n\times m}$，使得

$$\boldsymbol{AGA}=\boldsymbol{A},\ \boldsymbol{GAG}=\boldsymbol{G}$$

同时成立，则称 $\boldsymbol{G}$ 为 $\boldsymbol{A}$ 的自反广义逆矩阵，记为 $\boldsymbol{G}=\boldsymbol{A}_r^-$.

例 1 设 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r)\in\mathbf{U}_r^{m\times r}$, 即 $\boldsymbol{\alpha}_i^H\boldsymbol{\alpha}_j=\begin{cases}\mathbf{1},\ i=j\\ \mathbf{0},\ i\neq j\end{cases}\quad(i,j=1,2,\cdots,r)$, 则 $\boldsymbol{A}^H$ 为 $\boldsymbol{A}$ 的自反广义逆.

例 2 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$ 是行满秩矩阵，则 $\boldsymbol{A}$ 的右逆矩阵 $\boldsymbol{A}_R^{-1}$ 为 $\boldsymbol{A}$ 的一个自反广义逆.

定理 1 任何矩阵都有自反广义逆矩阵.

证明： (1) 若 $\boldsymbol{A}=\boldsymbol{O}$，则 $\boldsymbol{A}_r^-=\boldsymbol{O}$，显然成立.

(2) 对 $\boldsymbol{A}\neq\boldsymbol{O}$，不妨设 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}$，则存在可逆矩阵 $\boldsymbol{P}$, $\boldsymbol{Q}$，使得

$$\boldsymbol{A}=\boldsymbol{P}\begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\boldsymbol{Q},$$

取

$$\boldsymbol{G}=\boldsymbol{Q}^{-1}\begin{bmatrix}\boldsymbol{E}_r & \boldsymbol{X}\\ \boldsymbol{Y} & \boldsymbol{YX}\end{bmatrix}\boldsymbol{P}^{-1},$$

直接验证即可.

定理 2 设 $\boldsymbol{X},\boldsymbol{Y}\in\mathbf{C}^{n\times m}$ 均为 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$ 的广义逆矩阵，则

$$\boldsymbol{Z}=\boldsymbol{XAY}$$

是 $\boldsymbol{A}$ 的自反广义逆矩阵.

证明：因为 $\boldsymbol{X}$, $\boldsymbol{Y}$ 是 $\boldsymbol{A}$ 的广义逆矩阵，则 $\boldsymbol{AXA}=\boldsymbol{A}$, $\boldsymbol{AYA}=\boldsymbol{A}$，进而有

$$\boldsymbol{AZA}=\underline{\boldsymbol{AXA}}\boldsymbol{YA}=\boldsymbol{AYA}=\boldsymbol{A},$$

和

$$\boldsymbol{ZAZ}=\boldsymbol{X}\underline{\boldsymbol{AYA}}\boldsymbol{XAY}=\boldsymbol{X}\underline{\boldsymbol{AXA}}\boldsymbol{Y}=\boldsymbol{XAY}=\boldsymbol{Z},$$

得证.

定理 3 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，$\boldsymbol{A}^-$ 是 $\boldsymbol{A}$ 的广义逆矩阵，则 $\boldsymbol{A}^-$ 是 $\boldsymbol{A}$ 的自反广义逆矩阵的充要条件是

$$\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{A}^-).$$

证明： **(必要性)** 因为 $\boldsymbol{A}^-$ 是 $\boldsymbol{A}$ 的自反广义逆矩阵，即有

$$\boldsymbol{AA}^-\boldsymbol{A}=\boldsymbol{A},\ \boldsymbol{A}^-\boldsymbol{AA}^-=\boldsymbol{A}^-,$$

因此，

$$\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{AA}^-\boldsymbol{A})\le\mathrm{rank}(\boldsymbol{A}^-)\le\mathrm{rank}(\boldsymbol{A}^-\boldsymbol{AA}^-)\le\mathrm{rank}(\boldsymbol{A}),$$

即

$$\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{A}^-).$$

(充分性) 因为 $\boldsymbol{AA}^-\boldsymbol{A}=\boldsymbol{A}$，且 $\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{A}^-)$，因此

$$\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{A}^-\boldsymbol{A})=\mathrm{rank}(\boldsymbol{A}^-).$$

又由 $\mathbf{R}(\boldsymbol{A}^-\boldsymbol{A})\subset\mathbf{R}(\boldsymbol{A}^-)$ ，得 $\mathbf{R}(\boldsymbol{A}^-\boldsymbol{A})=\mathbf{R}(\boldsymbol{A}^-)$，故存在 $\boldsymbol{X}\in\mathbf{C}^{n\times m}$，使得 $\boldsymbol{A}^-=\boldsymbol{A}^-\boldsymbol{AX}$.

因为

$$\boldsymbol{A}=\boldsymbol{AA}^-\boldsymbol{A}=\boldsymbol{AA}^-\boldsymbol{AXA}=\boldsymbol{AXA},$$

即 $\boldsymbol{X}$ 为 $\boldsymbol{A}$ 的广义逆矩阵，由定理 2 知，$\boldsymbol{A}^-$ 为 $\boldsymbol{A}$ 的自反广义逆矩阵.

定理 4 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，$\boldsymbol{X}\in\mathbf{C}^{n\times m}$，则下列任意两个等式成立都可推导得第三个等式成立.

(1) $\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{X})$;

(2) $\boldsymbol{AXA}=\boldsymbol{A}$;

(3) $\boldsymbol{XAX}=\boldsymbol{X}$.

定理 5 设 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，则

$$\boldsymbol{X}=\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-}\boldsymbol{A}^{H},\ \boldsymbol{Y}=\boldsymbol{A}^{H}\left(\boldsymbol{A}\boldsymbol{A}^{H}\right)^{-}$$

都是 $\boldsymbol{A}$ 的自反广义逆矩阵.

证明: 因为 $\operatorname{rank}(\boldsymbol{A}^{H})=\operatorname{rank}(\boldsymbol{A}^{H}\boldsymbol{A})$, $\mathrm{R}(\boldsymbol{A}^{H}\boldsymbol{A})\subset\mathrm{R}(\boldsymbol{A}^{H})$，则有 $\mathrm{R}(\boldsymbol{A}^{H})=\mathrm{R}(\boldsymbol{A}^{H}\boldsymbol{A})$，于是存在矩阵 $\boldsymbol{D}\in\mathbf{C}^{n\times m}$，使得

$$\boldsymbol{A}^{H}=\boldsymbol{A}^{H}\boldsymbol{A}\boldsymbol{D},$$

即 $\boldsymbol{A}=\boldsymbol{D}^{H}\boldsymbol{A}^{H}\boldsymbol{A}$.

因为

$$\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}=\boldsymbol{A}\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-}\boldsymbol{A}^{H}\boldsymbol{A}=\boldsymbol{D}^{H}\boldsymbol{A}^{H}\boldsymbol{A}\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-}\boldsymbol{A}^{H}\boldsymbol{A}=\boldsymbol{D}^{H}\boldsymbol{A}^{H}\boldsymbol{A}=\boldsymbol{A},$$

所以 $\boldsymbol{X}$ 是 $\boldsymbol{A}$ 的广义逆矩阵.

又因

$$\operatorname{rank}(\boldsymbol{A})\leq\operatorname{rank}(\boldsymbol{X})=\operatorname{rank}\left(\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-}\boldsymbol{A}^{H}\right)\leq\operatorname{rank}(\boldsymbol{A}^{H})=\operatorname{rank}(\boldsymbol{A}),$$

即 $\operatorname{rank}(\boldsymbol{A})=\operatorname{rank}(\boldsymbol{X})$，由定理 3 知，$\boldsymbol{X}$ 是 $\boldsymbol{A}$ 的自反广义逆矩阵.

定理 6 设 $\boldsymbol{A}\boldsymbol{A}_r^{-}$ 和 $\boldsymbol{A}_r^{-}\boldsymbol{A}$ 都是幂等矩阵.

证明： 直接验证即可.

$$\left(\boldsymbol{A}\boldsymbol{A}_r^{-}\right)^{2}=\boldsymbol{A}\boldsymbol{A}_r^{-}\boldsymbol{A}\boldsymbol{A}_r^{-}=\boldsymbol{A}\boldsymbol{A}_r^{-},$$

即 $\boldsymbol{A}\boldsymbol{A}_r^{-}$ 是幂等矩阵.

同理，可证 $\boldsymbol{A}_r^{-}\boldsymbol{A}$ 是幂等矩阵.

§5 $M-P$ 广义逆矩阵 A^+

程光辉

2020 年 5 月 22 日

定义 1 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，如果存在矩阵 $\boldsymbol{G}\in\mathbf{C}^{n\times m}$，使得

$$\boldsymbol{AGA}=\boldsymbol{A},\ \boldsymbol{GAG}=\boldsymbol{G},\ (\boldsymbol{GA})^{H}=\boldsymbol{GA},\ (\boldsymbol{AG})^{H}=\boldsymbol{AG},$$

同时成立，则称 $\boldsymbol{G}$ 为 $\boldsymbol{A}$ 的 $\boldsymbol{M-P}$ 广义逆矩阵，记为 $\boldsymbol{G}=\boldsymbol{A}^{+}$.

定理 1 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，$\boldsymbol{A}=\boldsymbol{BD}$ 是 $\boldsymbol{A}$ 的最大秩分解，则

$$\boldsymbol{G}=\boldsymbol{D}^{H}\left(\boldsymbol{DD}^{H}\right)^{-1}\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H}$$

是 $\boldsymbol{A}$ 的广义逆矩阵 $\boldsymbol{A}^{+}$.

证明： (1) 若 $\mathbf{rank}(\boldsymbol{A})=\mathbf{0}$，即 $\boldsymbol{A}=\boldsymbol{O}$，则 $\boldsymbol{A}^{+}=\boldsymbol{O}$.

(2) 若 $\mathbf{rank}(\boldsymbol{A})>\mathbf{0}$,则存在最大秩分解 $\boldsymbol{A}=\boldsymbol{BD}$. 又因 $\mathbf{rank}(\boldsymbol{B}^{H}\boldsymbol{B})=\mathbf{rank}(\boldsymbol{DD}^{H})=\mathbf{rank}(\boldsymbol{A})$，所以 $\boldsymbol{B}^{H}\boldsymbol{B}$，$\boldsymbol{DD}^{H}$ 都是可逆的. 下面直接验证.

$$\begin{aligned}\boldsymbol{AGA}&=\boldsymbol{BDD}^{H}\left(\boldsymbol{DD}^{H}\right)^{-1}\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H}\boldsymbol{BD}\\&=\boldsymbol{BD}\\&=\boldsymbol{A}.\end{aligned}$$

$$\begin{aligned}\boldsymbol{GAG}&=\boldsymbol{D}^{H}\left(\boldsymbol{DD}^{H}\right)^{-1}\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H}\boldsymbol{BDD}^{H}\left(\boldsymbol{DD}^{H}\right)^{-1}\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H}\\&=\boldsymbol{D}^{H}\left(\boldsymbol{DD}^{H}\right)^{-1}\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H}\\&=\boldsymbol{G}.\end{aligned}$$

$$(\boldsymbol{GA})^{H}=\boldsymbol{A}^{H}\boldsymbol{G}^{H}$$

$$
\begin{aligned}
&= \boldsymbol{D}^H \boldsymbol{B}^H \boldsymbol{B}\left[\left(\boldsymbol{B}^H \boldsymbol{B}\right)^{-1}\right]^H \left[\left(\boldsymbol{D}\boldsymbol{D}^H\right)^{-1}\right]^H \boldsymbol{D} \\
&= \boldsymbol{D}^H \boldsymbol{B}^H \boldsymbol{B}\left(\boldsymbol{B}^H \boldsymbol{B}\right)^{-1} \left(\boldsymbol{D}\boldsymbol{D}^H\right)^{-1} \boldsymbol{D} \\
&= \boldsymbol{D}^H \left(\boldsymbol{D}\boldsymbol{D}^H\right)^{-1} \boldsymbol{D} \\
&= \boldsymbol{D}^H \left(\boldsymbol{D}\boldsymbol{D}^H\right)^{-1} \left(\boldsymbol{B}^H \boldsymbol{B}\right)^{-1} \boldsymbol{B}^H \boldsymbol{B}\boldsymbol{D} \\
&= \boldsymbol{G}\boldsymbol{A}.
\end{aligned}
$$

$$
\begin{aligned}
(\boldsymbol{A}\boldsymbol{G})^H &= \boldsymbol{G}^H \boldsymbol{A}^H \\
&= \boldsymbol{B}\left[\left(\boldsymbol{B}^H \boldsymbol{B}\right)^{-1}\right]^H \left[\left(\boldsymbol{D}\boldsymbol{D}^H\right)^{-1}\right]^H \boldsymbol{D}\boldsymbol{D}^H\boldsymbol{B}^H \\
&= \boldsymbol{B}\left(\boldsymbol{B}^H \boldsymbol{B}\right)^{-1} \left(\boldsymbol{D}\boldsymbol{D}^H\right)^{-1} \boldsymbol{D}\boldsymbol{D}^H\boldsymbol{B}^H \\
&= \boldsymbol{B}\left(\boldsymbol{B}^H \boldsymbol{B}\right)^{-1} \boldsymbol{B}^H \\
&= \boldsymbol{B}\boldsymbol{D}\boldsymbol{D}^H \left(\boldsymbol{D}\boldsymbol{D}^H\right)^{-1} \left(\boldsymbol{B}^H \boldsymbol{B}\right)^{-1} \boldsymbol{B}^H \\
&= \boldsymbol{A}\boldsymbol{G}.
\end{aligned}
$$

综上，$\boldsymbol{G}$ 是 $\boldsymbol{A}$ 的 $\boldsymbol{M}-\boldsymbol{P}$ 广义逆矩阵 $\boldsymbol{A}^+$.

定理 2 设 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，则 $\boldsymbol{A}^+$ 是唯一的.

证明： 令 $\boldsymbol{A}_1^+$，$\boldsymbol{A}_2^+$ 都是 $\boldsymbol{A}$ 的 $\boldsymbol{M}-\boldsymbol{P}$ 广义逆，则

$$
\begin{aligned}
\boldsymbol{A}_1^+ &= \boldsymbol{A}_1^+\boldsymbol{A}\boldsymbol{A}_1^+ \\
&= \boldsymbol{A}_1^+(\boldsymbol{A}\boldsymbol{A}_2^+\boldsymbol{A})\boldsymbol{A}_1^+ \\
&= \boldsymbol{A}_1^+(\boldsymbol{A}\boldsymbol{A}_2^+)(\boldsymbol{A}\boldsymbol{A}_1^+) \\
&= \boldsymbol{A}_1^+(\boldsymbol{A}\boldsymbol{A}_2^+)^H(\boldsymbol{A}\boldsymbol{A}_1^+)^H \\
&= \boldsymbol{A}_1^+\left(\boldsymbol{A}_2^+\right)^H\boldsymbol{A}^H\left(\boldsymbol{A}_1^+\right)^H\boldsymbol{A}^H \\
&= \boldsymbol{A}_1^+(\boldsymbol{A}_2^+)^H\left(\boldsymbol{A}\boldsymbol{A}_1^+\boldsymbol{A}\right)^H \\
&= \boldsymbol{A}_1^+(\boldsymbol{A}_2^+)^H(\boldsymbol{A})^H \\
&= \boldsymbol{A}_1^+(\boldsymbol{A}\boldsymbol{A}_2^+)^H \\
&= \boldsymbol{A}_1^+\boldsymbol{A}\boldsymbol{A}_2^+ \quad \text{(考虑左侧结合，重复上述步骤.)} \\
&= \boldsymbol{A}_2^+\boldsymbol{A}\boldsymbol{A}_2^+ \\
&= \boldsymbol{A},
\end{aligned}
$$

得证.

定理 3 设 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，则

(1) $\left(\boldsymbol{A}^{+}\right)^{+}=\boldsymbol{A}$;

(2) $\left(\boldsymbol{A}^{T}\right)^{+}=\left(\boldsymbol{A}^{+}\right)^{T}$, $\left(\boldsymbol{A}^{H}\right)^{+}=\left(\boldsymbol{A}^{+}\right)^{H}$;

(3) $\boldsymbol{A}^{+}=\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{+}\boldsymbol{A}^{H}=\boldsymbol{A}^{H}\left(\boldsymbol{A}\boldsymbol{A}^{H}\right)^{+}$;

(4) $\mathbf{R}(\boldsymbol{A}^{+})=\mathbf{R}(\boldsymbol{A}^{H})$;

(5) $\boldsymbol{A}\boldsymbol{A}^{+}=\mathbf{P}_{\mathbf{R}(\boldsymbol{A})}$, $\boldsymbol{A}^{+}\boldsymbol{A}=\mathbf{P}_{\mathbf{R}(\boldsymbol{A}^{H})}$;

(6) $\mathbf{R}(\boldsymbol{A})=\mathbf{R}(\boldsymbol{A}^{H})$ 的充要条件是 $\boldsymbol{A}\boldsymbol{A}^{+}=\boldsymbol{A}^{+}\boldsymbol{A}$.

证明： (1) 显然成立；

(2) 设 $\boldsymbol{A}=\boldsymbol{B}\boldsymbol{D}$ 是矩阵 $\boldsymbol{A}$ 的最大秩分解，则 $\boldsymbol{A}^{T}=\boldsymbol{D}^{T}\boldsymbol{B}^{T}$ 是矩阵 $\boldsymbol{A}^{T}$ 的最大秩分解. 于是有

$$\begin{aligned}
\left(\boldsymbol{A}^{+}\right)^{T}&=\left[\boldsymbol{D}^{H}\left(\boldsymbol{D}\boldsymbol{D}^{H}\right)^{-1}\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H}\right]^{T}\\
&=(\boldsymbol{B}^{H})^{T}\left[\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\right]^{T}\left[\left(\boldsymbol{D}\boldsymbol{D}^{H}\right)^{-1}\right]^{T}(\boldsymbol{D}^{H})^{T}\\
&=(\boldsymbol{B}^{T})^{H}\left[\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{T}\right]^{-1}\left[\left(\boldsymbol{D}\boldsymbol{D}^{H}\right)^{T}\right]^{-1}(\boldsymbol{D}^{T})^{H}\\
&=(\boldsymbol{B}^{T})^{H}\left[\boldsymbol{B}^{T}(\boldsymbol{B}^{T})^{H}\right]^{-1}\left[\boldsymbol{D}^{T}(\boldsymbol{D}^{T})^{H}\right]^{-1}(\boldsymbol{D}^{T})^{H}\\
&=(\boldsymbol{A}^{T})^{+}.
\end{aligned}$$

类似可证明 $\left(\boldsymbol{A}^{H}\right)^{+}=\left(\boldsymbol{A}^{+}\right)^{H}$.

(3) 设 $\boldsymbol{A}=\boldsymbol{B}\boldsymbol{D}$ 是矩阵 $\boldsymbol{A}$ 的最大秩分解，则 $\boldsymbol{A}^{H}\boldsymbol{A}=\boldsymbol{D}^{H}\boldsymbol{B}^{H}\boldsymbol{B}\boldsymbol{D}=\boldsymbol{B}_1\boldsymbol{D}_1$ 是矩阵 $\boldsymbol{A}^{H}\boldsymbol{A}$ 的最大秩分解，其中 $\boldsymbol{B}_1=\boldsymbol{D}^{H}$ 列满秩，$\boldsymbol{D}_1=\boldsymbol{B}^{H}\boldsymbol{B}\boldsymbol{D}$ 行满秩. 进而

$$\begin{aligned}
\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{+}\boldsymbol{A}^{H}&=\left[\boldsymbol{D}_1^{H}(\boldsymbol{D}_1\boldsymbol{D}_1^{H})^{-1}(\boldsymbol{B}_1^{H}\boldsymbol{B}_1)^{-1}\boldsymbol{B}_1^{H}\right](\boldsymbol{B}\boldsymbol{D})^{H}\\
&=(\boldsymbol{B}^{H}\boldsymbol{B}\boldsymbol{D})^{H}[(\boldsymbol{B}^{H}\boldsymbol{B}\boldsymbol{D})(\boldsymbol{B}^{H}\boldsymbol{B}\boldsymbol{D})^{H}]^{-1}(\boldsymbol{D}\boldsymbol{D}^{H})^{-1}\boldsymbol{D}\boldsymbol{D}^{H}\boldsymbol{B}^{H}\\
&=\boldsymbol{D}^{H}\boldsymbol{B}^{H}\boldsymbol{B}(\boldsymbol{B}^{H}\boldsymbol{B}\boldsymbol{D}\boldsymbol{D}^{H}\boldsymbol{B}^{H}\boldsymbol{B})^{-1}\boldsymbol{B}^{H}\\
&=\boldsymbol{D}^{H}\boldsymbol{B}^{H}\boldsymbol{B}(\boldsymbol{B}^{H}\boldsymbol{B})^{-1}(\boldsymbol{D}\boldsymbol{D}^{H})^{-1}(\boldsymbol{B}^{H}\boldsymbol{B})^{-1}\boldsymbol{B}^{H}\\
&=\boldsymbol{D}^{H}(\boldsymbol{D}\boldsymbol{D}^{H})^{-1}(\boldsymbol{B}^{H}\boldsymbol{B})^{-1}\boldsymbol{B}^{H}\\
&=\boldsymbol{A}^{+}.
\end{aligned}$$

类似，可证另外一个等式.

(4) 因 $\boldsymbol{A}^{+}$ 是 $\boldsymbol{A}$ 的自反广义逆，则 $\mathbf{rank}(\boldsymbol{A}^{+})=\mathbf{rank}(\boldsymbol{A})=\mathbf{rank}(\boldsymbol{A}^{H})$. 又由 (3) 知，$\mathbf{R}(\boldsymbol{A}^{+})\subset\mathbf{R}(\boldsymbol{A}^{H})$，进而 $\mathbf{R}(\boldsymbol{A}^{+})=\mathbf{R}(\boldsymbol{A}^{H})$.

(5) 因 $\boldsymbol{A}^{+}$ 是 $\boldsymbol{A}$ 的自反广义逆，则 $\boldsymbol{A}\boldsymbol{A}^{+}$ 和 $\boldsymbol{A}^{+}\boldsymbol{A}$ 都是幂等矩阵，即

$$\boldsymbol{A}\boldsymbol{A}^{+}=\mathbf{P}_{\mathbf{R}(\boldsymbol{A}\boldsymbol{A}^{+})},\quad \boldsymbol{A}^{+}\boldsymbol{A}=\mathbf{P}_{\mathbf{R}(\boldsymbol{A}^{+}\boldsymbol{A})}.$$

又因 $\mathbf{R}(\boldsymbol{A}\boldsymbol{A}^+) = \mathbf{R}(\boldsymbol{A})$，$\mathbf{R}(\boldsymbol{A}^+\boldsymbol{A}) = \mathbf{R}(\boldsymbol{A}^+)$，$\mathbf{R}(\boldsymbol{A}^+) = \mathbf{R}(\boldsymbol{A}^H)$，故

$$\boldsymbol{A}\boldsymbol{A}^+ = \mathbf{P}_{\mathbf{R}(\boldsymbol{A})}, \quad \boldsymbol{A}^+\boldsymbol{A} = \mathbf{P}_{\mathbf{R}(\boldsymbol{A}^H)}.$$

(6) 充分性：因 $\boldsymbol{A}\boldsymbol{A}^+ = \boldsymbol{A}^+\boldsymbol{A}$，结合 (5) 的证明，则

$$\mathbf{R}(\boldsymbol{A}) = \mathbf{R}(\boldsymbol{A}\boldsymbol{A}^+) = \mathbf{R}(\boldsymbol{A}^+\boldsymbol{A}) = \mathbf{R}(\boldsymbol{A}^+) = \mathbf{R}(\boldsymbol{A}^H).$$

必要性：因 $\mathbf{R}(\boldsymbol{A}) = \mathbf{R}(\boldsymbol{A}^H)$，$\boldsymbol{A}\boldsymbol{A}^+ = \mathbf{P}_{\mathbf{R}(\boldsymbol{A})}$，$\boldsymbol{A}^+\boldsymbol{A} = \mathbf{P}_{\mathbf{R}(\boldsymbol{A}^H)}$，则

$$\boldsymbol{A}\boldsymbol{A}^+ = \boldsymbol{A}^+\boldsymbol{A}.$$

定理 4 设 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，则有

(1) $(\boldsymbol{A}^H\boldsymbol{A})^+ = \boldsymbol{A}^+(\boldsymbol{A}^H)^+$，$(\boldsymbol{A}\boldsymbol{A}^H)^+ = (\boldsymbol{A}^H)^+\boldsymbol{A}^+$;

(2) $(\boldsymbol{A}^H\boldsymbol{A})^+ = \boldsymbol{A}^+(\boldsymbol{A}\boldsymbol{A}^H)^+\boldsymbol{A} = \boldsymbol{A}^H(\boldsymbol{A}\boldsymbol{A}^H)^+(\boldsymbol{A}^H)^+$;

(3) $\boldsymbol{A}\boldsymbol{A}^+ = (\boldsymbol{A}\boldsymbol{A}^H)(\boldsymbol{A}\boldsymbol{A}^H)^+ = (\boldsymbol{A}\boldsymbol{A}^H)^+(\boldsymbol{A}\boldsymbol{A}^H)$;
$\boldsymbol{A}^+\boldsymbol{A} = (\boldsymbol{A}^H\boldsymbol{A})(\boldsymbol{A}^H\boldsymbol{A})^+ = (\boldsymbol{A}^H\boldsymbol{A})^+(\boldsymbol{A}^H\boldsymbol{A})$.

证明： (1) 设 $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{D}$ 是矩阵 $\boldsymbol{A}$ 的最大秩分解，则 $\boldsymbol{A}^H\boldsymbol{A} = \boldsymbol{D}^H\boldsymbol{B}^H\boldsymbol{B}\boldsymbol{D} = \boldsymbol{B}_1\boldsymbol{D}_1$ 是矩阵 $\boldsymbol{A}^H\boldsymbol{A}$ 的最大秩分解，其中 $\boldsymbol{B}_1 = \boldsymbol{D}^H$ 列满秩，$\boldsymbol{D}_1 = \boldsymbol{B}^H\boldsymbol{B}\boldsymbol{D}$ 行满秩. 于是有

$$\begin{aligned}
\left(\boldsymbol{A}^H\boldsymbol{A}\right)^+ &= \boldsymbol{D}_1^H(\boldsymbol{D}_1\boldsymbol{D}_1^H)^{-1}(\boldsymbol{B}_1^H\boldsymbol{B}_1)^{-1}\boldsymbol{B}_1^H \\
&= (\boldsymbol{B}^H\boldsymbol{B}\boldsymbol{D})^H[(\boldsymbol{B}^H\boldsymbol{B}\boldsymbol{D})(\boldsymbol{B}^H\boldsymbol{B}\boldsymbol{D})^H]^{-1}(\boldsymbol{D}\boldsymbol{D}^H)^{-1}\boldsymbol{D} \\
&= \boldsymbol{D}^H\boldsymbol{B}^H\boldsymbol{B}(\boldsymbol{B}^H\boldsymbol{B}\boldsymbol{D}\boldsymbol{D}^H\boldsymbol{B}^H\boldsymbol{B})^{-1}(\boldsymbol{D}\boldsymbol{D}^H)^{-1}\boldsymbol{D} \\
&= \boldsymbol{D}^H\boldsymbol{B}^H\boldsymbol{B}(\boldsymbol{B}^H\boldsymbol{B})^{-1}(\boldsymbol{D}\boldsymbol{D}^H)^{-1}(\boldsymbol{B}^H\boldsymbol{B})^{-1}(\boldsymbol{D}\boldsymbol{D}^H)^{-1}\boldsymbol{D} \\
&= \boldsymbol{D}^H(\boldsymbol{D}\boldsymbol{D}^H)^{-1}(\boldsymbol{B}^H\boldsymbol{B})^{-1}(\boldsymbol{D}\boldsymbol{D}^H)^{-1}\boldsymbol{D} \\
&= \left[\boldsymbol{D}^H(\boldsymbol{D}\boldsymbol{D}^H)^{-1}(\boldsymbol{B}^H\boldsymbol{B})^{-1}\boldsymbol{B}^H\right]\left[\boldsymbol{B}(\boldsymbol{B}^H\boldsymbol{B})^{-1}(\boldsymbol{D}\boldsymbol{D}^H)^{-1}\boldsymbol{D}\right] \\
&= \boldsymbol{A}^+(\boldsymbol{A}^H)^+.
\end{aligned}$$

类似，可证另外一个等式.

(2) 定理 3 知，$\left(\boldsymbol{A}^H\right)^+ = \left(\boldsymbol{A}^+\right)^H = \left[\boldsymbol{A}^H\left(\boldsymbol{A}\boldsymbol{A}^H\right)^+\right]^H = \left(\boldsymbol{A}\boldsymbol{A}^H\right)^+\boldsymbol{A}^H$. 又由 (1) 知，$(\boldsymbol{A}^H\boldsymbol{A})^+ = \boldsymbol{A}^+(\boldsymbol{A}^H)^+$，进而

$$(\boldsymbol{A}^H\boldsymbol{A})^+ = \boldsymbol{A}^+(\boldsymbol{A}^H)^+ = \boldsymbol{A}^+\left(\boldsymbol{A}\boldsymbol{A}^H\right)^+\boldsymbol{A}^H.$$

类似，可证另外一个等式.

(3) 由定理 3(3)，知

$$\boldsymbol{A}\boldsymbol{A}^+ = \boldsymbol{A}\left[\boldsymbol{A}^H\left(\boldsymbol{A}\boldsymbol{A}^H\right)^+\right] = \left(\boldsymbol{A}\boldsymbol{A}^H\right)\left(\boldsymbol{A}\boldsymbol{A}^H\right)^+.$$

又因为

$$\begin{aligned}
\left(\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}}\right)\left(\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}}\right)^{+} &= \left[\left(\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}}\right)\left(\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}}\right)^{+}\right]^{\boldsymbol{H}} \\
&= \left[\left(\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}}\right)^{+}\right]^{\boldsymbol{H}}\left(\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}}\right)^{\boldsymbol{H}} \\
&= \left(\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}}\right)^{+}\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}} \\
&= \boldsymbol{A}\boldsymbol{A}^{+}.
\end{aligned}$$

故有

$$\boldsymbol{A}\boldsymbol{A}^{+} = (\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}})(\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}})^{+} = (\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}})^{+}(\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}}).$$

类似，可证另外一个等式.

定理 5 设 $\boldsymbol{A} \in \mathbf{C}^{m\times l}$，$\boldsymbol{B} \in \mathbf{C}^{l\times n}$，则

$$(\boldsymbol{A}\boldsymbol{B})^{+} = \boldsymbol{B}^{+}\boldsymbol{A}^{+} \Leftrightarrow \mathrm{R}(\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{A}\boldsymbol{B}) \subset \mathrm{R}(\boldsymbol{B})\text{且}\mathrm{R}(\boldsymbol{B}\boldsymbol{B}^{\boldsymbol{H}}\boldsymbol{A}^{\boldsymbol{H}}) \subset \mathrm{R}(\boldsymbol{A}^{\boldsymbol{H}}).$$

§6 A^+ 的计算方法

程光辉

2019 年 12 月 17 日

1 最大秩分解

引理 1 设 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，则

(1) 如果 $\boldsymbol{A}$ 是行满秩矩阵，则 $\boldsymbol{A}^+ = \boldsymbol{A}^H(\boldsymbol{A}\boldsymbol{A}^H)^{-1}$;

(2) 如果 $\boldsymbol{A}$ 是列满秩矩阵，则 $\boldsymbol{A}^+ = (\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H$.

证明： (1) 因为 $\boldsymbol{A} \in \mathbf{C}_m^{m\times n}$，则 $\boldsymbol{A} = \boldsymbol{E}_m\boldsymbol{A}$ 是矩阵 $\boldsymbol{A}$ 的最大秩分解，于是有

$$\boldsymbol{A}^+ = \boldsymbol{A}^H\left(\boldsymbol{A}\boldsymbol{A}^H\right)^{-1}\left(\boldsymbol{E}_m^H\boldsymbol{E}_m\right)^{-1}\boldsymbol{E}_m^H = \boldsymbol{A}^H\left(\boldsymbol{A}\boldsymbol{A}^H\right)^{-1}.$$

(2) 类似 (1) 的证明.

定理 1 设 $\boldsymbol{A} \in \mathbf{C}_r^{m\times n}$，$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{D}$ 是 $\boldsymbol{A}$ 的最大秩分解，则

$$\boldsymbol{A}^+ = \boldsymbol{D}^+\boldsymbol{B}^+.$$

证明： 因为 $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{D}$ 是 $\boldsymbol{A}$ 的最大秩分解，由引理 1 可得

$$\boldsymbol{A}^+ = \boldsymbol{D}^H\left(\boldsymbol{D}\boldsymbol{D}^H\right)^{-1}\left(\boldsymbol{B}^H\boldsymbol{B}\right)^{-1}\boldsymbol{B}^H = \boldsymbol{D}^+\boldsymbol{B}^+.$$

例 1 设矩阵 $\boldsymbol{A}$ 为

$$\boldsymbol{A} = \begin{pmatrix} 1 & 3 & 2 & 1 & 4 \\ 2 & 6 & 1 & 0 & 7 \\ 3 & 9 & 3 & 1 & 11 \end{pmatrix},$$

求 $\boldsymbol{A}$ 的 $\boldsymbol{M-P}$ 广义逆矩阵 $\boldsymbol{A}^+$.

解： (1) 求 $\boldsymbol{A}$ 的最大秩分解 $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{D}$，

$$\boldsymbol{B} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 3 \end{pmatrix}, \quad \boldsymbol{D} = \begin{pmatrix} 1 & 3 & 0 & -\frac{1}{3} & \frac{10}{3} \\ 0 & 0 & 1 & \frac{2}{3} & \frac{1}{3} \end{pmatrix}.$$

(2) 计算 $\boldsymbol{B}^{+}$ 和 $\boldsymbol{D}^{+}$，

$$\boldsymbol{B}^{+}=\left(\boldsymbol{B}^{H}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{H}=-\frac{1}{9}\begin{pmatrix}-4 & 5 & 1\\ 5 & -1 & 1\end{pmatrix},$$

$$\boldsymbol{D}^{+}=\boldsymbol{D}^{H}\left(\boldsymbol{D}\boldsymbol{D}^{H}\right)^{-1}=\frac{9}{290}\begin{pmatrix}1 & 0\\ 3 & 0\\ 0 & 1\\ -\frac{1}{3} & \frac{2}{3}\\ \frac{10}{3} & \frac{1}{3}\end{pmatrix}.$$

(3) 计算 $\boldsymbol{A}^{+}=\boldsymbol{D}^{+}\boldsymbol{B}^{+}$，

$$\boldsymbol{A}^{+}=\frac{3}{290}\begin{pmatrix}-32 & 34 & 2\\ -96 & 102 & 6\\ 329 & -268 & 61\\ 230 & -190 & 40\\ 3 & 24 & 27\end{pmatrix}.$$

例 2 举例说明下列结论不成立：

(1) $(\boldsymbol{AB})^{+}=\boldsymbol{B}^{+}\boldsymbol{A}^{+}$.

(2) $(\boldsymbol{A}^{k})^{+}=(\boldsymbol{A}^{+})^{k}$, 其中 $\boldsymbol{k}$ 是正整数.

(3) 若 $\boldsymbol{P}$, $\boldsymbol{Q}$ 为可逆矩阵，$(\boldsymbol{PAQ})^{+}=\boldsymbol{Q}^{-1}\boldsymbol{A}^{+}\boldsymbol{P}^{-1}$.

解： (1) 令 $\boldsymbol{A}=\begin{pmatrix}1 & 1\end{pmatrix}$，$\boldsymbol{B}=\begin{pmatrix}1\\ 0\end{pmatrix}$，则有 $\boldsymbol{AB}=(1)$，$(\boldsymbol{AB})^{+}=(1)$.

因为 $\boldsymbol{A}$ 行满秩，则 $\boldsymbol{A}^{+}=\boldsymbol{A}^{H}(\boldsymbol{A}\boldsymbol{A}^{H})^{-1}=\frac{1}{2}\begin{pmatrix}1\\ 1\end{pmatrix}$. 因为 $\boldsymbol{B}$ 列满秩，则 $\boldsymbol{B}^{+}=(\boldsymbol{B}^{H}\boldsymbol{B})^{-1}\boldsymbol{B}^{H}=\begin{pmatrix}1 & 0\end{pmatrix}$.

于是有 $\boldsymbol{B}^{+}\boldsymbol{A}^{+}=(\frac{1}{2})\neq(\boldsymbol{AB})^{+}$.

(2) 令 $\boldsymbol{A}=\begin{pmatrix}1 & -1\\ 0 & 0\end{pmatrix}$，则 $\boldsymbol{A}$ 为幂等矩阵.

$\boldsymbol{A}=\begin{pmatrix}1\\ 0\end{pmatrix}\begin{pmatrix}1 & -1\end{pmatrix}$ 是其最大秩分解，则 $\boldsymbol{A}^{+}=\frac{1}{2}\begin{pmatrix}1 & 0\\ -1 & 0\end{pmatrix}$.

于是有

$$(\boldsymbol{A}^{2})^{+}=\boldsymbol{A}^{+}=\frac{1}{2}\begin{pmatrix}1 & 0\\ -1 & 0\end{pmatrix},$$

$$(\boldsymbol{A}^+)^2 = \frac{1}{4}\begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix},$$

则有 $(\boldsymbol{A}^2)^+ \neq (\boldsymbol{A}^+)^2$.

(3) 令 $\boldsymbol{A} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$，$\boldsymbol{P} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$，$\boldsymbol{Q} = (1)$，则有

$$\boldsymbol{A}^+ = (\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H = \frac{1}{2}\begin{pmatrix} 1 & 1 \end{pmatrix},$$

$$(\boldsymbol{PAQ})^+ = \begin{pmatrix} 2 \\ 1 \end{pmatrix}^+ = \frac{1}{5}\begin{pmatrix} 2 & 1 \end{pmatrix},$$

$$\boldsymbol{Q}^{-1}\boldsymbol{A}^+\boldsymbol{P}^{-1} = (1)\frac{1}{2}\begin{pmatrix} 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 \end{pmatrix},$$

显然有 $(\boldsymbol{PAQ})^+ \neq \boldsymbol{Q}^{-1}\boldsymbol{A}^+\boldsymbol{P}^{-1}$.

2 奇异值分解法

定理 2 设 $\boldsymbol{A} \in \mathbf{C}_r^{m\times n}$ 的奇异值分解为

$$\boldsymbol{A} = \boldsymbol{U}\begin{pmatrix} \boldsymbol{D}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}\boldsymbol{V}^H = \boldsymbol{UDV}^H,$$

则有

(1) $\boldsymbol{A}^+ = \boldsymbol{VD}^+\boldsymbol{U}^H$*;*

(2) $\|\boldsymbol{A}^+\|_F^2 = \sum\limits_{i=1}^{r}\frac{1}{\sigma_i^2}$*;*

(3) $\|\boldsymbol{A}^+\|_2 = \frac{1}{\min\limits_{1\leq i\leq r}\{\sigma_i\}}$.

证明： (1) 直接验证. (2) 和 (3) 利用范数的酉不变性.

定理 3 设 $\boldsymbol{A} \in \mathbf{C}^{m\times n}$，$\boldsymbol{\lambda}_i$ $(i = 1, 2, \cdots, r)$ 是 $\boldsymbol{AA}^H$ 的 r 个非零特征值，$\boldsymbol{\alpha}_i$ $(i = 1, 2, \cdots, r)$ 是 $\boldsymbol{AA}^H$ 对应于 $\boldsymbol{\lambda}_i$ 单位正交的特征向量，记 $\boldsymbol{\Delta}_r = \mathbf{diag}(\boldsymbol{\lambda}_1, \cdots, \boldsymbol{\lambda}_r)$, $\boldsymbol{U}_1 = (\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_r)$，则有

$$\boldsymbol{A}^+ = \boldsymbol{A}^H\boldsymbol{U}_1\boldsymbol{\Delta}_r^{-1}\boldsymbol{U}_1^H.$$

证明： 设 $\boldsymbol{A} \in \mathbf{C}_r^{m\times n}$ 的奇异值分解为

$$\boldsymbol{A} = \boldsymbol{U}\begin{pmatrix} \boldsymbol{D}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}\boldsymbol{V}^H,$$

则有

$$\begin{aligned}\boldsymbol{AA^H} &= \boldsymbol{U}\begin{pmatrix}\boldsymbol{D_r} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{V^HV}\begin{pmatrix}\boldsymbol{D_r^H} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{U^H}\\ &= \boldsymbol{U}\begin{pmatrix}\boldsymbol{\Delta_r} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\boldsymbol{U^H},\end{aligned}$$

于是有

$$\begin{aligned}(\boldsymbol{AA^H})^+ &= \boldsymbol{U}\begin{pmatrix}\boldsymbol{\Delta_r} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}^+\boldsymbol{U^H}\\ &= \begin{pmatrix}\boldsymbol{U_1} & \boldsymbol{U_2}\end{pmatrix}\begin{pmatrix}\boldsymbol{\Delta_r^{-1}} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\begin{pmatrix}\boldsymbol{U_1^H}\\ \boldsymbol{U_2^H}\end{pmatrix}\\ &= \boldsymbol{U_1\Delta_r^{-1}U_1^H}.\end{aligned}$$

即

$$\boldsymbol{A^+} = \boldsymbol{A^H(AA^H)^+} = \boldsymbol{A^HU_1\Delta_r^{-1}U_1^H}.$$

例 3 设矩阵 $\boldsymbol{A}$ 为

$$\boldsymbol{A} = \begin{pmatrix}-1 & 0 & 1\\ 2 & 0 & -2\end{pmatrix},$$

求 $\boldsymbol{A}$ 的 $\boldsymbol{M-P}$ 广义逆矩阵 $\boldsymbol{A^+}$.

解： 求 $\boldsymbol{AA^H}$ 的特征值和非零特征值对应的单位正交特征向量，

$$\boldsymbol{AA^H} = \begin{pmatrix}2 & -4\\ -4 & 8\end{pmatrix},$$

则特征值为 $\boldsymbol{\lambda_1 = 10}$，$\boldsymbol{\lambda_2 = 0}$. 进而 $\boldsymbol{\lambda_1 = 10}$ 对应的单位特征向量为 $\boldsymbol{\alpha_1} = \begin{pmatrix}\frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}}\end{pmatrix}^T$，即 $\boldsymbol{U_1 = \alpha_1}$.

由定理 3，得

$$\boldsymbol{A^+} = \boldsymbol{A^HU_1\Delta_r^{-1}U_1^H} = \begin{pmatrix}-1 & 2\\ 0 & 0\\ 1 & -2\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{5}}\\ -\frac{2}{\sqrt{5}}\end{pmatrix}\left(\frac{1}{10}\right)\begin{pmatrix}\frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}}\end{pmatrix}.$$

§7 广义逆矩阵的应用

程光辉

2020 年 1 月 1 日

1 矩阵方程的通解

定理 1 设 $A \in \mathbf{C}^{m\times n}$，$B \in \mathbf{C}^{p\times q}$，$D \in \mathbf{C}^{m\times q}$，则矩阵方程

$$AXB = D$$

有解的充要条件是存在 A^- 和 B^-，使得

$$AA^-DB^-B = D$$

成立. 在有解的条件下，矩阵方程 $AXB = D$ 的通解为

$$X = A^-DB^- + Y - A^-AYBB^-,\quad \forall Y \in \mathbf{C}^{n\times p}.$$

证明： (**必要性**) 设 X 为 $AXB = D$ 的解，因为 $AA^-A = A$，$BB^-B = B$，则有

$$D = AXB = AA^-AXBB^-B = AA^-DB^-B.$$

(**充分性**) 因为 $D = AA^-DB^-B$，故 $X = A^-DB^-$ 是 $AXB = D$ 的解.

下面证明在有解情况下的通解. 因为

$$\begin{aligned}A(A^-DB^- + Y - A^-AYBB^-)B &= AA^-DB^-B + AYB - AA^-AYBB^-B\\ &= D + AYB - AYB\\ &= D,\end{aligned}$$

因此，$X = A^-DB^- + Y - A^-AYBB^-$ 是 $AXB = D$ 的解.

设 G 是 $AXB = D$ 的任一解，即 $AGB = D$，进而

$$G = A^-DB^- + G - A^-DB^- = A^-DB^- + G - A^-AGBB^-,$$

故矩阵方程 $AXB = D$ 的通解为

$$X = A^-DB^- + Y - A^-AYBB^-,\quad \forall Y \in \mathbf{C}^{n\times p}.$$

推论 1 设 $A \in \mathbf{C}^{m\times n}$，$D \in \mathbf{C}^{m\times p}$，则矩阵方程 $AX = D$ 有解的充要条件是存在 A^-，使得

$$AA^-D = D$$

成立. 在有解的条件下，矩阵方程 $AX = D$ 的通解为

$$X = A^-D + Y - A^-AY, \quad \forall Y \in \mathbf{C}^{n\times p}.$$

推论 2 设 $B \in \mathbf{C}^{m\times n}$，$D \in \mathbf{C}^{p\times n}$，则矩阵方程 $XB = D$ 有解的充要条件是存在 B^-，使得

$$DB^-B = D$$

成立. 在有解的条件下，矩阵方程 $XB = D$ 的通解为

$$X = DB^- + Y - YBB^-, \quad \forall Y \in \mathbf{C}^{p\times m}.$$

推论 3 设 $A \in \mathbf{C}^{m\times n}$，$b \in \mathbf{C}^{m}$，则线性方程组 $Ax = b$ 有解的充要条件是存在 A^-，使得

$$AA^-b = b$$

成立. 在有解的条件下，线性方程组 $Ax = b$ 的通解为

$$x = A^-b + (E_n - A^-A)u, \quad \forall u \in \mathbf{C}^{n}.$$

2 相容方程的最小范数解

定义 1 方程组 $Ax = b$ 有解，则称此方程组为相容方程组；否则，称为不相容方程组.

定义 2 设方程组 $Ax = b$ 有解，将所有的解中范数最小的解称为最小范数解.

令 $A\{1,3\} = \{G | AGA = A, \ (GA)^H = GA\}.$

定理 2 设 $A \in \mathbf{C}^{m\times n}$，$D \in A\{1,3\}$，则 Db 是相容方程组 $Ax = b$ 的最小范数解，并且方程组的最小范数解唯一.

证明： 设 x_0 是 $Ax = b$ 的任意解，$D \in A\{1,3\}$，则有 $ADA = A$，进而

$$b = Ax_0 = ADAx_0 = ADb,$$

于是有 Db 是 $Ax = b$ 的解. 根据定理 1 的推论 3 知，$Ax = b$ 的通解为

$$x = Db + (E - DA)u, \quad \forall u \in \mathbf{C}^{n}.$$

对任意解 $\boldsymbol{x_0}$ 有，

$$\begin{aligned}\|\boldsymbol{x_0}\|_2^2 &= \|\boldsymbol{Db}+(\boldsymbol{E}-\boldsymbol{DA})\boldsymbol{u}\|_2^2 \\ &= \left(\boldsymbol{Db}+(\boldsymbol{E}-\boldsymbol{DA})\boldsymbol{u}\right)^{\boldsymbol{H}}\left(\boldsymbol{Db}+(\boldsymbol{E}-\boldsymbol{DA})\boldsymbol{u}\right) \\ &= \|\boldsymbol{Db}\|_2^2+\|(\boldsymbol{E}-\boldsymbol{DA})\boldsymbol{u}\|_2^2+(\boldsymbol{Db})^{\boldsymbol{H}}(\boldsymbol{E}-\boldsymbol{DA})\boldsymbol{u}+\boldsymbol{u}^{\boldsymbol{H}}(\boldsymbol{E}-\boldsymbol{DA})^{\boldsymbol{H}}\boldsymbol{Db}. \quad (1)\end{aligned}$$

又由 $\boldsymbol{Ax_0}=\boldsymbol{b}$ 得

$$\begin{aligned}(\boldsymbol{Db})^{\boldsymbol{H}}(\boldsymbol{E}-\boldsymbol{DA})\boldsymbol{u} &= (\boldsymbol{DAx_0})^{\boldsymbol{H}}(\boldsymbol{E}-\boldsymbol{DA})\boldsymbol{u} \\ &= \boldsymbol{x_0^H}(\boldsymbol{DA})^{\boldsymbol{H}}(\boldsymbol{E}-\boldsymbol{DA})\boldsymbol{u} \\ &= \boldsymbol{x_0^H DA}(\boldsymbol{E}-\boldsymbol{DA})\boldsymbol{u} \\ &= \boldsymbol{x_0^H}(\boldsymbol{DA}-\boldsymbol{DADA})\boldsymbol{u} \\ &= \boldsymbol{0},\end{aligned}$$

所以由 (1) 式可得

$$\|\boldsymbol{x_0}\|_2^2=\|\boldsymbol{Db}\|_2^2+\|(\boldsymbol{E}-\boldsymbol{DA})\boldsymbol{u}\|_2^2\geq\|\boldsymbol{Db}\|_2^2,$$

则 $\boldsymbol{Db}$ 是相容方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小范数解.

唯一性显然.

定理 3 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，$\boldsymbol{D}\in\mathbf{C}^{n\times m}$，$\forall\boldsymbol{b}\in\mathbf{C}^{m}$，则 $\boldsymbol{Db}$ 是相容方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小范数解，则必有 $\boldsymbol{D}\in\boldsymbol{A}\{1,3\}$.

证明：设 $\boldsymbol{A}=(\boldsymbol{\alpha_1},\cdots,\boldsymbol{\alpha_n})$，则 $\boldsymbol{\alpha_i}\in\mathbf{R}(\boldsymbol{A})$. 设 $\boldsymbol{G}\in\boldsymbol{A}\{1,3\}$，则 $\boldsymbol{G\alpha_i}$ 是 $\boldsymbol{Ax}=\boldsymbol{\alpha_i}$ 的最小范数解. 由定理 2 知最小范数解是唯一的，则有

$$\boldsymbol{D\alpha_i}=\boldsymbol{G\alpha_i},\quad\forall i=1,2,\cdots,n,$$

即 $\boldsymbol{DA}=\boldsymbol{GA}$，进而 $\boldsymbol{ADA}=\boldsymbol{AGA}=\boldsymbol{A}$，$(\boldsymbol{DA})^{\boldsymbol{H}}=(\boldsymbol{GA})^{\boldsymbol{H}}=\boldsymbol{GA}=\boldsymbol{DA}$，于是有 $\boldsymbol{D}\in\boldsymbol{A}\{1,3\}$.

3 不相容方程组的解

如果 $\boldsymbol{Ax}=\boldsymbol{b}$ 是不相容方程组，令 $\boldsymbol{f}(\boldsymbol{x})=\|\boldsymbol{Ax}-\boldsymbol{b}\|_2^2$，存在 $\boldsymbol{x_0}$ 使得 $\boldsymbol{f}(\boldsymbol{x_0})$ 最小，称 $\boldsymbol{x_0}$ 为方程组的最小二乘解，这种问题称为最小二乘解问题.

令 $\boldsymbol{A}\{1,4\}=\{\boldsymbol{G}|\boldsymbol{AGA}=\boldsymbol{A},\ (\boldsymbol{AG})^{\boldsymbol{H}}=\boldsymbol{AG}\}$.

定理 4 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，$\boldsymbol{G}\in\boldsymbol{A}\{1,4\}$，则 $\boldsymbol{Gb}$ 是不相容方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小二乘解.

证明：因为

$$\begin{aligned}\|\boldsymbol{Ax}-\boldsymbol{b}\|_2^2 &= \|\boldsymbol{Ax}-\boldsymbol{AGb}+\boldsymbol{AGb}-\boldsymbol{b}\|_2^2 \\ &= (\boldsymbol{Ax}-\boldsymbol{AGb}+\boldsymbol{AGb}-\boldsymbol{b})^{\boldsymbol{H}}(\boldsymbol{Ax}-\boldsymbol{AGb}+\boldsymbol{AGb}-\boldsymbol{b}) \\ &= \|\boldsymbol{Ax}-\boldsymbol{AGb}\|_2^2+\|\boldsymbol{AGb}-\boldsymbol{b}\|_2^2 \\ &+ (\boldsymbol{Ax}-\boldsymbol{AGb})^{\boldsymbol{H}}(\boldsymbol{AGb}-\boldsymbol{b})+(\boldsymbol{AGb}-\boldsymbol{b})^{\boldsymbol{H}}(\boldsymbol{Ax}-\boldsymbol{AGb}). \end{aligned} \tag{2}$$

又由 $(\boldsymbol{AG})^{\boldsymbol{H}}=\boldsymbol{AG}$，有

$$\begin{aligned}(\boldsymbol{AGb}-\boldsymbol{b})^{\boldsymbol{H}}(\boldsymbol{Ax}-\boldsymbol{AGb}) &= (\boldsymbol{b}^{\boldsymbol{H}}(\boldsymbol{AG})^{\boldsymbol{H}}-\boldsymbol{b}^{\boldsymbol{H}})(\boldsymbol{Ax}-\boldsymbol{AGb}) \\ &= (\boldsymbol{b}^{\boldsymbol{H}}\boldsymbol{AG}-\boldsymbol{b}^{\boldsymbol{H}})(\boldsymbol{Ax}-\boldsymbol{AGb}) \\ &= \boldsymbol{b}^{\boldsymbol{H}}\boldsymbol{AGAx}-\boldsymbol{b}^{\boldsymbol{H}}\boldsymbol{Ax}-\boldsymbol{b}^{\boldsymbol{H}}\boldsymbol{AGAGb}+\boldsymbol{b}^{\boldsymbol{H}}\boldsymbol{AGb} \\ &= \boldsymbol{0}, \end{aligned}$$

因此，由式 (2) 可得

$$\|\boldsymbol{Ax}-\boldsymbol{b}\|_2^2=\|\boldsymbol{Ax}-\boldsymbol{AGb}\|_2^2+\|\boldsymbol{AGb}-\boldsymbol{b}\|_2^2\geq\|\boldsymbol{AGb}-\boldsymbol{b}\|_2^2,$$

即 $\boldsymbol{Gb}$ 是不相容方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小二乘解.

引理 1 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，$\boldsymbol{x}$ 是不相容方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小二乘解的充要条件是 $\boldsymbol{Ax}=\boldsymbol{AGb}$，$\forall\boldsymbol{G}\in\boldsymbol{A}\{1,4\}$.

证明：(**必要性**) 对 $\forall\boldsymbol{G}\in\boldsymbol{A}\{1,4\}$，$\forall\boldsymbol{x}\in\mathbf{C}^n$，由定理 4 证明知，

$$\|\boldsymbol{Ax}-\boldsymbol{b}\|_2^2=\|\boldsymbol{Ax}-\boldsymbol{AGb}\|_2^2+\|\boldsymbol{AGb}-\boldsymbol{b}\|_2^2=\|\boldsymbol{AGb}-\boldsymbol{b}\|_2^2,$$

则 $\boldsymbol{Ax}=\boldsymbol{AGb}$.

(**充分性**) 若 $\boldsymbol{Ax}=\boldsymbol{AGb}$，则 $\|\boldsymbol{Ax}-\boldsymbol{b}\|_2^2=\|\boldsymbol{AGb}-\boldsymbol{b}\|_2^2$, 即 $\boldsymbol{x}$ 是不相容方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小二乘解.

定理 5 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，$\boldsymbol{G}\in\boldsymbol{A}\{1,4\}$，不相容方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小二乘解的通解是 $\boldsymbol{x}=\boldsymbol{Gb}+(\boldsymbol{E}-\boldsymbol{A}^-\boldsymbol{A})\boldsymbol{u}$, $\forall\boldsymbol{u}\in\mathbf{C}^n$.

证明：令 $\boldsymbol{x}$ 是不相容方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小二乘解，则有 $\boldsymbol{Ax}=\boldsymbol{AGb}$，即 $\boldsymbol{x}$ 是相容方程组 $\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{Gb})=\boldsymbol{0}$ 的解. 再由推论 3，知通解为

$$\boldsymbol{x}-\boldsymbol{Gb}=(\boldsymbol{E}-\boldsymbol{A}^-\boldsymbol{A})\boldsymbol{u},\ \forall\boldsymbol{u}\in\mathbf{C}^n,$$

即

$$\boldsymbol{x}=\boldsymbol{Gb}+(\boldsymbol{E}-\boldsymbol{A}^-\boldsymbol{A})\boldsymbol{u},\ \forall\boldsymbol{u}\in\mathbf{C}^n.$$

定义 3 设 $\boldsymbol{x}_0$ 是不相容方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小二乘解，若对方程组的任意最小二乘解 $\bar{\boldsymbol{x}}$，均有

$$\|\boldsymbol{x}_0\|_2 \le \|\bar{\boldsymbol{x}}\|_2,$$

则称 $\boldsymbol{x}_0$ 是不相容方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最佳最小二乘解，简称最佳逼近解.

定理 6 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，不相容方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最佳逼近解是 $\boldsymbol{x}=\boldsymbol{A}^{+}\boldsymbol{b}$.

例 1 用广义逆矩阵方法判断线性方程组

$$\begin{cases} 2x_1+4x_2+x_3+x_4=3 \\ x_1+2x_2-x_3+2x_4=0 \\ -x_1-2x_2-2x_3+x_4=3 \end{cases}$$

是否有解？如果有解，求通解和最小范数解；如果无解，求最小二乘解和最佳逼近解.

解： 令系数矩阵和右端向量为

$$\boldsymbol{A}=\begin{pmatrix} 2 & 4 & 1 & 1 \\ 1 & 2 & -1 & 2 \\ -1 & -2 & -2 & 1 \end{pmatrix},\quad \boldsymbol{b}=\begin{pmatrix} 3 \\ 0 \\ 3 \end{pmatrix},$$

则线性方程组即为 $\boldsymbol{Ax}=\boldsymbol{b}$.

因为 $\boldsymbol{A}^{+}\in\boldsymbol{A}\{1,3\}$ 且 $\boldsymbol{A}^{+}\in\boldsymbol{A}\{1,4\}$，再由定理 2 和定理 4 知，可利用 $\boldsymbol{A}^{+}\boldsymbol{b}$ 判断方程组解的类型，所以首先计算 $\boldsymbol{A}^{+}\boldsymbol{b}$.

矩阵 $\boldsymbol{A}$ 的最大秩分解为 $\boldsymbol{A}=\boldsymbol{BD}$，其中

$$\boldsymbol{B}=\begin{pmatrix} 2 & 1 \\ 1 & -1 \\ -1 & -2 \end{pmatrix},\quad \boldsymbol{D}=\begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

则

$$\boldsymbol{A}^{+}=\boldsymbol{D}^{H}(\boldsymbol{D}\boldsymbol{D}^{H})^{-1}(\boldsymbol{B}^{H}\boldsymbol{B})^{-1}\boldsymbol{B}^{H}=\frac{1}{33}\begin{pmatrix} 2 & 1 & -1 \\ 4 & 2 & -2 \\ 1 & -5 & -6 \\ 1 & 6 & 5 \end{pmatrix}.$$

又因为

$$\boldsymbol{A}\boldsymbol{A}^{+}\boldsymbol{b}=\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}\ne\boldsymbol{b},$$

故 $\boldsymbol{Ax}=\boldsymbol{b}$ 是不相容方程组.

于是有不相容方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的最小二乘解的通解为

$$\boldsymbol{x}=\boldsymbol{A}^{+}\boldsymbol{b}+(\boldsymbol{E}-\boldsymbol{A}^{+}\boldsymbol{A})\boldsymbol{u}=\frac{\mathbf{1}}{\mathbf{11}}\begin{pmatrix}\mathbf{1}\\\mathbf{2}\\\mathbf{-5}\\\mathbf{6}\end{pmatrix}+(\boldsymbol{E}-\boldsymbol{A}^{+}\boldsymbol{A})\boldsymbol{u},\quad\forall\boldsymbol{u}\in\mathbf{C}^{4},$$

最佳逼近解为

$$\boldsymbol{x}=\boldsymbol{A}^{+}\boldsymbol{b}=\frac{\mathbf{1}}{\mathbf{11}}\begin{pmatrix}\mathbf{1}\\\mathbf{2}\\\mathbf{-5}\\\mathbf{6}\end{pmatrix}.$$