代

a

矩阵分析

(最后编辑于 2016/04/21 - 19:13:00)

http://leptc.github.io/lenote

精 Horn. Matrix Analysis (2). Cambridge
　└ 中译: 杨奇. 矩阵分析. 机械工业出版社

参 张贤达. 矩阵分析与应用（第二版）. 清华大学出版社

(矩阵分析属研究生课程, 上面推荐的两本书都相当厚, 我也只是读了其中最基础的部分, 此外本笔记还收纳了来自线代, 高代, 信号, 信息等学科的参考书的内容)

符号约定

斜体为标量, 粗斜小写为矢量, 粗斜大写为矩阵

相关笔记

相关算法见〈数值分析〉 概率应用见〈概率论〉

矩阵术语

matrix 矩阵 $A=\begin{bmatrix} a_{11} & \cdots & a_{1n}\\ & \ddots & \\ a_{m1} & \cdots & a_{mn}\end{bmatrix}$ a_{ii} 称为 主对角线 main diagonal ，方阵 square matrix 从右上到左下称为 副 / 次 / 交叉对角线 counter / secondary / cross / antidiagonal

非主对角元全为零称为 对角阵 diagonal matrix diag$(a_{11}\sim a_{nn})$

$a_{ii}=1$ 的对角方阵称为 单位阵 $I_n=\begin{bmatrix}e_1 & \dots & e_n\end{bmatrix}$, 所有元素为 0 称为 零矩阵 $O_{m\times n}$, 零矢量 记为 $\mathbf{0}$

$j-i>p$ 和 $i-j>q$ 的阵元都为零称为 带型矩阵 banded matrix , p,q 称为 上 / 下带宽

下/上带宽为 0 的方阵称为 上 / 下三角阵 upper / lower triangular matrix , 对角元为 0/1 称为 严格 / 单位上 / 下三角 strictly / unit

性质 上/下三角阵之积, 幂, 逆仍为上/下三角阵, 幂的对角元为 a_{ii}^k , 本征值就是对角元

至少有一边带宽为 1 的方阵称为 上 / 下海森堡阵 , 上下带宽均为 1 的方阵称为 三对角阵

block matrix 分块矩阵 以维数匹配的矩阵为矩阵元 **例** 任意不满秩的矩阵若分块成 $\begin{bmatrix}A_{r\times r} & A_2\\ A_3 & A_4\end{bmatrix}$

则 $A_4=A_3A_r^{-1}A_2$ 「因为线性相关, $\exists B$ 使 $A_3=BA_r$, $A_4=BA_2$ 」

一元运算

记 $A=[a_{ij}]_{m\times n}$, 转置 transpose $A^{\mathsf{T}}\equiv[a_{ji}]_{n\times m}$ 复共轭 complex conjugate $A^*\equiv[a_{ij}^*]_{m\times n}$ 厄米共轭 Hermitian conjugate $A^\dagger\equiv[a_{ji}^*]_{n\times m}=(A^*)^{\mathsf{T}}=(A^{\mathsf{T}})^*$

性质 转置, 共轭, 厄米均可和求逆运算换序 $(A^\dagger)^{-1}=(A^{-1})^\dagger\equiv A^{-\dagger}$ 「$(AA^{-1})^{\mathsf{T}}=I\to(A^{-1})^{\mathsf{T}}A^{\mathsf{T}}=I$」

性质 转置, 厄米, 逆对矩阵乘积均要换序分配 $(AB)^{-1}=B^{-1}A^{-1}$

公式 $(A^{-1}+I)^{-1}=[(I+A)A^{-1}]^{-1}=A(A+I)^{-1}\to(A^{-1}+B^{-1})^{-1}=A(A+B)^{-1}B=B(A+B)^{-1}A$

Sherman-Morrison 矩阵求逆引理 (1949) $(A+xy^\dagger)^{-1}=A^{-1}-A^{-1}xy^\dagger A^{-1}/(1+y^\dagger A^{-1}x)$ 「$(I+B)^{-1}=I-B+B^2-\dots$」

矩阵之和求逆公式 $(A+UBV)^{-1}=A^{-1}-A^{-1}U(I+BVA^{-1}U)^{-1}BVA^{-1}$ (形式不唯一)

Woodbury 伍德伯里公式 $(A-UV)^{-1}=A^{-1}-A^{-1}UC^{-1}VA^{-1}$, 其中 $C\equiv I+VA^{-1}U$ 称为 容量矩阵 capacitance matrix

分块矩阵求逆公式 (形式不唯一) , 其中 $X\equiv(D-VA^{-1}U)^{-1}$

$$\begin{bmatrix}A & U\\ V & D\end{bmatrix}^{-1}\xlongequal{\text{若 }A\text{ 可逆}}\begin{bmatrix}A^{-1}+A^{-1}UXVA^{-1} & -A^{-1}UX\\ -XVA^{-1} & X\end{bmatrix}\xlongequal{\text{若 }A,D\text{ 可逆}}\begin{bmatrix}(A-UD^{-1}V)^{-1} & -(V-DU^{-1}A)^{-1}\\ (U-AV^{-1}D)^{-1} & X\end{bmatrix}$$

$\to$ 分块三角阵求逆公式 (若相应出现的逆存在) $\begin{bmatrix}A & O\\ V & D\end{bmatrix}^{-1}=\begin{bmatrix}A^{-1} & O\\ -D^{-1}VA^{-1} & D^{-1}\end{bmatrix}$, $\begin{bmatrix}A & U\\ V & O\end{bmatrix}^{-1}=\begin{bmatrix}O & -V^{-1}\\ U^{-1} & -U^{-1}AV^{-1}\end{bmatrix}$

特殊矩阵

若 $A^2=A$ 称为 幂等阵 idempotent **性质** $A^k=A$, 都可对角化, 除单位阵外都奇异

设 rank $A=r$, 则 A 的本征值为 r 个 1 和 $n-r$ 个 0 , tr $A=r$ **定理** 若 A,B 幂等且对易, 则 AB 幂等

推论 若 A 幂等, 则 $A^\dagger$ 幂等, $I-A$ 幂等 ($A-I$ 则不一定) 秩为 $n-r$, $A(I-A)=(I-A)A=O$

若 $A^{\mathsf{T}}A=A$ 称为 对称幂等阵 **性质** $A=A^{\mathsf{T}}=A^2$, 除单位阵外都半正定

若 $A^2=I$ 称为 幂单阵 unipotent $\Leftrightarrow A^{-1}=A$ 对合阵 involutory / involution , $A^{-1}=A^*$ 称为 共轭对合阵 coninvolutory

定理 A 幂单 $\Leftrightarrow\frac{1}{2}(A+I)$ 幂等 $\Leftrightarrow(I-A)(I+A)=O$

若 $A^k=O$ 称为 幂零阵 nilpotent , 使之成立的最小 $k\in\mathbb{N}_+$ 称为 幂零指标 index of nilpotence **定理** 幂零阵 $\Leftrightarrow$ 所有本征值都是 0

定理 若幂零, 则 k 不需要超过其维数, 即 $A^{\dim V}$ 必 $=O$

permutation 置换阵 每行每列都只有 1 个 1 的方阵 (重排了的单位阵) EA 对行重排, AE 对列重排 〈线代〉

性质 是正交阵, 同阶置换阵之积仍为置换阵 **定理** 对于置换阵, 必存在 k 使 $E^k=I$ 〈群论〉

reversal / exchange / backward identity 反演 / 交换 / 反向单位阵 $J=\begin{bmatrix}e_n & e_{n-1} & \dots & e_1\end{bmatrix}$ **性质** 幂单, 对称, 汉克尔阵, 副对角阵

basic circulant permutation

基本轮换阵 $\boldsymbol{C}_n=[\boldsymbol{e}_n\ \boldsymbol{e}_1\ \dots\ \boldsymbol{e}_{n-1}]$ 性质 常对角, 不幂单, 不对称, 轮换阵 $\boldsymbol{C}_n^k$ 之间可对易

vectorization commutation

记 $\mathrm{vec}(\boldsymbol{A})$ 将 $\boldsymbol{A}$ 按列 矢量化 , 记 换位阵 实现矩阵元转置 $\boldsymbol{K}_{mn}\mathrm{vec}(\boldsymbol{A}_{m\times n})=\mathrm{vec}(\boldsymbol{A}^\top)$

则 $\boldsymbol{K}_{mn}$ 是 $mn\times mn$ 的置换阵, $\mathrm{rank}(\boldsymbol{K}_{mn})=1+\gcd(m-1,n-1)$ 例 $\boldsymbol{K}_{1n}=\boldsymbol{K}_{n1}=\boldsymbol{I}_n$

性质 $\boldsymbol{K}_{mn}^\top=\boldsymbol{K}_{mn}^{-1}=\boldsymbol{K}_{nm}\to\boldsymbol{K}_{nn}$ 幂单, 对称, 本征值为 $\frac{1}{2}n(n+1)$ 个 1 和 $\frac{1}{2}n(n-1)$ 个 -1

diagonal-constant / Toeplitz

常对角阵 / 特普利茨阵 对角方向元素相同 $a_{ij}=a_{j-i}$

$$\begin{bmatrix} a_0 & a_1 & \dots & a_n \\ a_{-1} & a_0 & \dots & a_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{-n} & a_{1-n} & \dots & a_0 \end{bmatrix}$$

性质 副对称, 可表示为反演阵乘汉克尔阵

(求解常对角线性方程组有递推算法, 用 FFT 可降计算复杂度)

backward shift forward shift

后移 $\boldsymbol{B}=[\boldsymbol{0}\ \boldsymbol{e}_1\ \dots\ \boldsymbol{e}_{n-1}]$ 前移 $\boldsymbol{F}=[\boldsymbol{e}_2\ \dots\ \boldsymbol{e}_{n-1}\ \boldsymbol{0}]=\boldsymbol{B}^\top$, 常对角阵 $\Leftrightarrow$ 可表为 $\sum_1^n a_{-k}\boldsymbol{F}^k+\sum_0^n a_k\boldsymbol{B}^k$

Hankel

汉克尔阵 副对角方向元素相同 $a_{ij}=a_{i+j-2}$ 性质 对称

$$\begin{bmatrix} a_0 & a_1 & \dots & a_n \\ a_1 & a_2 & \dots & a_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_n & a_{n+1} & \dots & a_{2n} \end{bmatrix}$$

Hilbert

希尔伯特阵 (1894) $a_{ij}=\dfrac{1}{i+j-1}$ 性质 正定, 病态, $|\boldsymbol{A}|=\dfrac{c_n^4}{c_{2n}}$, $c_n=\prod_{i=1}^{n-1}i^{n-i}=\prod_{i=1}^{n-1}i!$

Cauchy

柯西阵 $a_{ij}=\dfrac{1}{a_i+b_j}$, $\forall(a_i+b_j)\neq0$, (Schechter 1959) $|\boldsymbol{A}|=\dfrac{\prod_{1\leqslant i<j\leqslant n}(a_j-a_i)(b_j-b_i)}{\prod_{i=1}^n\prod_{j=1}^n(a_i+b_j)}\neq0$ 总可逆

对称厄米

symmetric anti / skew symmetric Hermitian / self-adjoint anti-Hermite

对称阵 $\boldsymbol{A}^\top=\boldsymbol{A}$ 反 / 斜对称阵 $\boldsymbol{A}^\top=-\boldsymbol{A}$ 厄米 / 自伴阵 (1855) $\boldsymbol{A}^\dagger=\boldsymbol{A}$ 反厄米阵 $\boldsymbol{A}=-\boldsymbol{A}^\dagger$

推论 若 $\boldsymbol{A},\boldsymbol{B}$ 厄米, 则 $\boldsymbol{A}+\boldsymbol{B}$ 厄米, 若对易则 $\boldsymbol{AB}$ 厄米, $\boldsymbol{A}^k$ $(k\in\mathbb{N}_+)$ 均厄米, 若可逆则 $\boldsymbol{A}^{-1}$ 厄米

若 $\boldsymbol{A},\boldsymbol{B}$ 反厄米, 则 $\boldsymbol{A}^{-1}$, $\boldsymbol{A}+\boldsymbol{B}$ 反厄米, $\boldsymbol{A}^k$ 奇数幂反厄米, 偶数幂厄米

若 $\boldsymbol{A}$, $\boldsymbol{B}$ 厄米/反厄米, 则 $\alpha\boldsymbol{A}+\beta\boldsymbol{B}$ $(\alpha,\beta\in\mathbb{R})$ 厄米/反厄米, $\mathbf{i}\boldsymbol{A}$ 反厄米/厄米

若 $\boldsymbol{A}$, $\boldsymbol{B}$ 厄米, 则 $\boldsymbol{AB}+\boldsymbol{BA}$ 厄米, $\boldsymbol{AB}-\boldsymbol{BA}$ 反厄米

定理 对任意矩阵, $\boldsymbol{AA}^\dagger$, $\boldsymbol{A}^\dagger\boldsymbol{A}$ 厄米, 对任意方阵, $\boldsymbol{A}+\boldsymbol{A}^\dagger$ 厄米, $\boldsymbol{A}-\boldsymbol{A}^\dagger$ 反厄米

任意方阵均可分解为厄米和反厄米部分 $\boldsymbol{A}=\frac{1}{2}(\boldsymbol{A}+\boldsymbol{A}^\dagger)+\frac{1}{2}(\boldsymbol{A}-\boldsymbol{A}^\dagger)$ 定理 正规阵 $\Leftrightarrow$ 两部分对易

性质 $\boldsymbol{A}^\top$ 和 $\boldsymbol{A}$ 本征值相同, 本征矢不同, $\boldsymbol{A}^\dagger$ 和 $\boldsymbol{A}^*$ 同理

$\boldsymbol{A}$ 和 $\boldsymbol{A}^*$ 的本征值共轭 $\to$ 厄米阵的本征值为实数, 反厄米的本征值为纯虚数 (可类比为实部和虚部)

「设若尔当块 $\boldsymbol{A}=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}$, 则 $\boldsymbol{JAJ}=\begin{bmatrix}0&0&0\\1&0&0\\0&1&0\end{bmatrix}$, $\boldsymbol{JA}=\begin{bmatrix}0&0&0\\0&0&1\\0&1&0\end{bmatrix}$, $\boldsymbol{AJ}=\begin{bmatrix}0&1&0\\1&0&0\\0&0&0\end{bmatrix}$

记 $\boldsymbol{S}=\frac{1}{\sqrt{2}}(\boldsymbol{I}+\mathbf{i}\boldsymbol{J})$, 则 $\boldsymbol{S}$ 对称幺正, $\boldsymbol{S}^{-1}=\boldsymbol{S}^*$ $\boldsymbol{SAS}^{-1}=\frac{1}{2}(\boldsymbol{A}+\boldsymbol{JAJ})+\frac{\mathbf{i}}{2}(\boldsymbol{JA}-\boldsymbol{AJ})$ 」

定理 任意方阵可相似到复对称阵 (实对称则不一定) 「$\lambda\boldsymbol{I}+\boldsymbol{SAS}^{-1}$」 $\to\boldsymbol{A}\sim\boldsymbol{A}^\top$

alternating matrix

交错阵 满足 $\boldsymbol{A}^\top=-\boldsymbol{A}$ 且对角元均为零, 在特征 $\neq2$ 的域上交错 $\Leftrightarrow$ 斜称

在 二元域 $\mathbb{F}_2=\{0,1\}$ 上, $-1=1$, 单位阵斜称不交错, 交错 $\Rightarrow$ 斜称 $\Leftrightarrow$ 对称

persymmetric perhermitian skew persymmetric

副对称阵 关于副对角线对称的方阵 $\boldsymbol{AJ}=\boldsymbol{JA}^\top$ 副厄米阵 $a_{ij}=a^*_{n+1-j,n+1-i}$ 斜副对称阵 $\boldsymbol{AJ}=-\boldsymbol{JA}^\top$

性质 副对称阵的转置, 逆, 之积, 仍副对称

centrosymmetric centrohermitian

中心对称阵 关于中央对称 $\boldsymbol{AJ}=\boldsymbol{JA}$ 中心厄米阵 $\boldsymbol{A}=\boldsymbol{JA}^*\boldsymbol{J}\Leftrightarrow a_{ij}=a^*_{n+1-i,n+1-j}$

$$\begin{bmatrix} a & x & y \\ z & b & z \\ y & x & a \end{bmatrix}$$

性质 中心对称阵的转置, 逆仍中心对称, 两个 斜中心对称 之积为中心对称

正交幺正

orthogonal unitary

正交阵 $\boldsymbol{QQ}^\top=\boldsymbol{Q}^\top\boldsymbol{Q}=\boldsymbol{I}\Leftrightarrow\boldsymbol{Q}^{-1}=\boldsymbol{Q}^\top$ 幺正 / 酉阵 $\boldsymbol{UU}^\dagger=\boldsymbol{U}^\dagger\boldsymbol{U}=\boldsymbol{I}\Leftrightarrow\boldsymbol{U}^{-1}=\boldsymbol{U}^\dagger$

幺正阵和复正交阵的交集为实正交阵 例 $\begin{bmatrix}\cosh t & \mathbf{i}\sinh t\\ -\mathbf{i}\sinh t & \cosh t\end{bmatrix}$ 是复正交阵, 仅当 $t=0$ 为幺正阵

(与幺正群不同, 复正交群不是有界集, 故不是紧集, 通常正交群指实正交阵组成的较小的紧群) 〈群论〉

性质 幺正矩阵必可逆, 本征值的模为 1 , 即位于单位圆上 $\lambda=\mathrm{e}^{\mathrm{i}\theta},\theta\in\mathbb{R}$

proper rotation

实正交阵 $\lambda=\pm1$, 取 $+1$ 则称 $\boxed{\text{真转动}}$, 复正交阵 $|A|=1$ 但 $|\lambda|$ 可以任意大

推论 $U^\top, U^*, U^\dagger, U^{-1}, U^k$ $(k\in\mathbb{N}_+)$ 均幺正

若 A,B 幺正, 则 $AB, A\oplus B, A\otimes B$ 均幺正 (没有加法)

定理 U 幺正 $\Leftrightarrow$ 行/列为标准正交组 $\Leftrightarrow$ 保内积 $\langle x,y\rangle=x^\dagger y=\langle Ux,Uy\rangle$ $\to$ 保长度, 夹角 (转动, 镜像)

$A\sim U \Rightarrow A^{-1}\sim A^\dagger$ (逆不成立, 例如 $A=\mathrm{diag}(2,\frac{1}{2})$)

skew orthogonal

$\boxed{\text{斜正交阵}}$ $A^{-1}=-A^\top \Leftrightarrow \pm\mathrm{i}A$ 是正交阵, 一般地 $A^{-1}=\mathrm{e}^{\mathrm{i}\theta}A^\top \Leftrightarrow \mathrm{e}^{\mathrm{i}\theta/2}A$ 是正交阵

semi-orthogonal para-unitary

非方阵 $Q_{m\times n}$ 若只满足 $QQ^\top=I_m$ 或 $Q^\top Q=I_n$ 称为 $\boxed{\text{半正交阵}}$, 复数域称为 $\boxed{\text{仿幺正阵}}$

(对称, 正交, 幂单, 满足两者则第三者也成立)

$\boxed{\text{对称幺正阵}}$ $\Leftrightarrow$ 可写成 $UU^\top$ $\Leftrightarrow$ 可写成 $Q\,\mathrm{diag}(\mathrm{e}^{\mathrm{i}\theta_{1\sim n}})Q^\top$ $\boxed{\text{厄米幺正阵}}$ $A=A^\dagger=A^{-1}$ **例** 泡利阵

Householder

$\boxed{\text{豪斯赫尔德阵}}$ $H=I-tee^\dagger$, 若 $e^\dagger e=1$ 则 $t=2$ **性质** 厄米幺正, 本征值为 $-1,1,1,\ldots$ $\to |H|=-1$

作用: 关于 $e^\perp$ 面做镜面反射「在补空间 $e^\perp$ 上不变, 在 e 子空间上反号」

正规对易

normal

方阵满足 $A^\dagger A=AA^\dagger$ 称为 $\boxed{\text{正规阵}}$ $\leftarrow$ 厄米, 反厄米/交错, 幺正阵均正规

(此外还可以是 $\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}$ (不限于二元域) , 注意复对称阵 $\neq$ 厄米阵, 可能不正规 $\begin{bmatrix}\mathrm{i} & \mathrm{i}\\ \mathrm{i} & 1\end{bmatrix}$)

性质 若 A 正规, $\Leftrightarrow A+\varphi I$ 正规, $\Rightarrow$ 多项式 $f(A)$ 正规, 两正规阵对易 $\Rightarrow$ 乘积正规

定理「类比复数 $z^*=\mathrm{e}^{\mathrm{i}\theta}z \to \exists U$ 使 $A^*=AU=UA$」 A 正规 $\Leftrightarrow \exists U$ 使 $A^\dagger=AU$

A 正规 $\Leftrightarrow A=U^\dagger\Lambda U \Leftrightarrow A=UH$, 其中 U 幺正 H 厄米且两者对易

将复对称阵分解为实虚部 $A=B+\mathrm{i}C$, 则 A 正规 $\Leftrightarrow B,C$ 对易 $\Leftrightarrow A,A^\dagger$ 对易 $\Leftrightarrow AA^\dagger$ 为实矩阵

A 正规 $\Leftrightarrow \forall x, \|Ax\|=\|A^\dagger x\|$, A 和 $A^\dagger$ 的本征矢相同

commutator anti-commutator

$\boxed{\text{对易式 / 换位子}}$ $[A,B]\equiv AB-BA$ $\boxed{\text{反对易式}}$ $\{A,B\}\equiv AB+BA$

推论 $\mathrm{tr}([A,B])=0$, $[A,B]^\dagger=[B^\dagger,A^\dagger]$ **例** 若 $[A,[A,B]]=O$, 则 $[A,B]$ 必幂零

性质 交换反对称, 线性 $[A+B,C]=[A,C]+[B,C]$, 雅可比恒等式〈 群论 〉

乘积展开 $[AB,C]=A[B,C]+[A,C]B$, $[A,BC]=B[A,C]+[A,B]C$

commute anticommute

$AB=BA$ 称两矩阵 $\boxed{\text{对易 / 可交换}}$ $AB=-BA$ 称 $\boxed{\text{反对易}}$

若 $[A,B]$ 对易, 则 $[A,f(B)]$ 仍对易, 若 B 可逆则 $\Leftrightarrow [A,B^{-1}]$ 对易, 若还有 $[A^\dagger,B]$ 对易 $\Leftrightarrow A$ 正规

simultaneously diagonalizable

若 $S^{-1}AS, S^{-1}BS$ 均对角, 称 A,B $\boxed{\text{可同时对角化}}$ $\Rightarrow A,B$ 对易 (逆命题需前提: A,B 均可对角化)

commuting family

(意义: 在某个基下看, 有且只有所有的对角阵对易) $\to$ 可推广到 $\boxed{\text{交换族}}$

定理 若 A,B 对易, 则存在一个排列, 使 $A+B$ 的本征值为 $\lambda_i^{(A)}+\lambda_j^{(B)}$ (逆命题不成立)

同阶方阵 A,B 本征矢相同 (本征值不必相同) , 其中一者的本征值互异, 则 A,B 对易

通过合同同时对角化分以下情况, 设 A 可逆, $C=A^{-1}B$「证明用到 SVD」

① A,B 厄米, 则 A,B 可通过 $\dagger$ 合同同时对角化 $\Leftrightarrow C$ 可相似到实对角阵 $\Leftrightarrow A+\mathrm{i}B$ 和正规阵 $\dagger$ 合同

可通过幺正 $\dagger$ 合同同时对角化 $\Leftrightarrow C$ 可幺正相似到实对角阵 $\Leftrightarrow C$ 厄米 $\Leftrightarrow A,B$ 对易 $\Leftrightarrow AB$ 厄米

② A,B 对称, 则 A,B 可通过 $\top$ 合同同时对角化 $\Leftrightarrow C$ 可相似对角化

可通过幺正 $\top$ 合同同时对角化 $\Leftrightarrow C$ 可幺正相似对角化 $\Leftrightarrow C$ 正规

例 $\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$ 可幺正 $\top$ 合同同时对角化, 但通过 $\dagger$ 合同不行

相似合同

right eigenvector left eigenvector

$(A-\lambda I)v=0 \Leftrightarrow Av=\lambda v$ 称为 $\boxed{\text{右本征矢}}$, 而 $\boxed{\text{左本征矢}}$ $u^\dagger(A-\lambda I)=0^\top \Leftrightarrow u^\dagger A=\lambda u^\dagger \Leftrightarrow A^\dagger u=\lambda^* u$

定理 若 $\lambda_u \neq \lambda_v$, 则 $\boldsymbol{u},\boldsymbol{v}$ 正交「结合 $\boldsymbol{u}^\dagger \boldsymbol{A}\boldsymbol{v}=\lambda_u \boldsymbol{u}^\dagger \boldsymbol{v}=\lambda_v \boldsymbol{u}^\dagger \boldsymbol{v}$」若 $\lambda_u=\lambda_v$ 且 $\boldsymbol{u},\boldsymbol{v}$ 不正交, 则 $\boldsymbol{A}\sim\begin{bmatrix}\lambda & \boldsymbol{0}^\mathsf{T}\\ \boldsymbol{0} & \boldsymbol{B}_{(n-1)^2}\end{bmatrix}$

推论 厄米阵同一 λ 对应相同左右本征矢 $\boldsymbol{v}=\boldsymbol{u}$ $\to$ 厄米阵的不同本征值的本征矢正交

(复对称阵不同本征值的本征矢满足 $\boldsymbol{u}_1^\mathsf{T}\boldsymbol{u}_2=0$ 这并不代表正交, 复对称阵不一定可对角化)

〈 线代 〉「谱定理」上述性质可推广到正规阵, $\boldsymbol{A}\boldsymbol{v}=\lambda\boldsymbol{v} \Leftrightarrow \boldsymbol{v}^\dagger\boldsymbol{A}=\lambda\boldsymbol{v}^\dagger \Leftrightarrow \boldsymbol{A}^\dagger\boldsymbol{v}=\lambda^*\boldsymbol{v}$

满足 $\boldsymbol{B}=\boldsymbol{U}^\dagger\boldsymbol{A}\boldsymbol{U}$ 称为 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ **幺正相似** (既相似又合同) **舒尔三角化** (Schur triangularization) 复方阵可幺正相似到上三角阵

推论 幺正相似阵的 $\mathrm{tr}(\boldsymbol{A}^\dagger\boldsymbol{A})=\sum_{ij}|a_{ij}|^2$ 相等「 亦可从几何上保内积来证」

舒尔定理 $\mathrm{tr}(\boldsymbol{A}^2)=\sum|\lambda_i|^2=\sum_{ij}|a_{ij}|^2-\sum_{i<j}|r_{ij}|^2\leqslant\mathrm{tr}(\boldsymbol{A}^\dagger\boldsymbol{A})$, 仅当可对角化 $\Leftrightarrow$ 正规阵时取等号

Specht theorem

施佩希特定理 $\boldsymbol{A},\boldsymbol{B}$ 幺正相似 $\Leftrightarrow$ 对所有 **字** (word) 都有 $\mathrm{tr}\,W(\boldsymbol{A},\boldsymbol{A}^\dagger)=\mathrm{tr}\,W(\boldsymbol{B},\boldsymbol{B}^\dagger)$ 〈 群论 〉

Pearcy theorem

皮尔希定理 不必验证无穷多个字, 二维只需验证 a,a^2,ab , 三维只需验证 $a,a^2,ab,a^3,a^2b,a^2b^2,a^2b^2ab$

若存在可逆方阵 $\boldsymbol{S}$ 使同阶方阵 $\boldsymbol{B}=\boldsymbol{S}^{-*}\boldsymbol{A}\boldsymbol{S}$ 则称两矩阵 **合相似** (consimilar) , **酉合相似** $\Leftrightarrow$ $\boldsymbol{B}=\boldsymbol{U}^\mathsf{T}\boldsymbol{A}\boldsymbol{U}$

方阵可合三角化 $\Leftrightarrow$ 可酉合三角化 $\Leftrightarrow$ $\boldsymbol{A}\boldsymbol{A}^*$ 所有本征值非负实

$\boldsymbol{A}\boldsymbol{x}^*=\lambda\boldsymbol{x}$ 称为 **合本征矢** (coneigenvector) 和 **合本征值** (coneigenvalue) , 若 $\lambda\geqslant0$ 则 $\Leftrightarrow$ λ^2 是 $\boldsymbol{A}\boldsymbol{A}^*$ 的本征值

二元运算

矩阵的 **加法** $\boldsymbol{A}+\boldsymbol{B}\equiv[a_{ij}+b_{ij}]$ **数乘** $\varphi\boldsymbol{A}\equiv[\varphi a_{ij}]$ **性质** 见〈 线代 〉线性空间

性质 转置, 共轭, 厄米对矩阵加法均有分配律 $(\boldsymbol{A}+\boldsymbol{B})^\dagger=\boldsymbol{A}^\dagger+\boldsymbol{B}^\dagger$

对 $\boldsymbol{A}_{m\times n}$, $\boldsymbol{B}_{r\times s}$, 仅当 $n=r$ 时有 **矩阵乘法** $(\boldsymbol{A}\boldsymbol{B})_{ij}\equiv\sum_{k=1}^n a_{ik}b_{kj}$, $i=1\sim m$, $j=1\sim s$ 〈 线代 〉

直和

方阵的 **直和** (direct sum) $\boldsymbol{A}_{m\times m}\oplus\boldsymbol{B}_{n\times n}\equiv\begin{bmatrix}\boldsymbol{A} & \boldsymbol{O}_{m\times n}\\ \boldsymbol{O}_{n\times m} & \boldsymbol{B}\end{bmatrix}_{(m+n)\times(m+n)}$

性质 没有交换律, 有结合律, 数乘分配律 $\varphi(\boldsymbol{A}\oplus\boldsymbol{B})=\varphi\boldsymbol{A}\oplus\varphi\boldsymbol{B}$, 转置, 共轭, 厄米, 逆均有分配律

$\mathrm{tr}(\bigoplus\boldsymbol{A}_i)=\sum\mathrm{tr}\,\boldsymbol{A}_i$, $\mathrm{rank}(\bigoplus\boldsymbol{A}_i)=\sum\mathrm{rank}\,\boldsymbol{A}_i$, $\det(\bigoplus\boldsymbol{A}_i)=\prod\det\boldsymbol{A}_i$

$(\boldsymbol{A}\oplus\boldsymbol{B})^{-1}=\boldsymbol{A}^{-1}\oplus\boldsymbol{B}^{-1}\to(\boldsymbol{A}\oplus\boldsymbol{B})^\#=|\boldsymbol{B}|\boldsymbol{A}^\#\oplus|\boldsymbol{A}|\boldsymbol{B}^\#$

设 $\mathrm{eig}(\boldsymbol{A})=(\lambda_A,\boldsymbol{u}_A)$, $\mathrm{eig}(\boldsymbol{B})=(\lambda_B,\boldsymbol{u}_B)$, 则 $\mathrm{eig}(\boldsymbol{A}\oplus\boldsymbol{B})=(\{\lambda_A,\lambda_B\},\{\begin{bmatrix}\boldsymbol{u}_A\\ \boldsymbol{0}\end{bmatrix},\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{u}_B\end{bmatrix}\})$

例 $\boldsymbol{A},\boldsymbol{B}$ 同阶, $\boldsymbol{C},\boldsymbol{D}$ 同阶, 有 $(\boldsymbol{A}\pm\boldsymbol{B})\oplus(\boldsymbol{C}\pm\boldsymbol{D})=(\boldsymbol{A}\oplus\boldsymbol{C})\pm(\boldsymbol{B}\oplus\boldsymbol{D})$, $(\boldsymbol{A}\oplus\boldsymbol{C})(\boldsymbol{B}\oplus\boldsymbol{D})=\boldsymbol{A}\boldsymbol{B}\oplus\boldsymbol{C}\boldsymbol{D}$

定理 正交阵的直和仍为正交阵

直积

direct / tensor / Kronecker product

直积 / 张量积 / 右克罗内克积 $\boldsymbol{A}_{m\times n}\otimes\boldsymbol{B}_{p\times q}\equiv\begin{bmatrix}a_{11}\boldsymbol{B} & \dots & a_{1n}\boldsymbol{B}\\ & \ddots & \\ a_{m1}\boldsymbol{B} & \dots & a_{mn}\boldsymbol{B}\end{bmatrix}_{mp\times nq}$ **左克罗内克积** $\boldsymbol{A}\otimes_{\text{left}}\boldsymbol{B}\equiv\begin{bmatrix}\boldsymbol{A}b_{11} & \dots & \boldsymbol{A}b_{1n}\\ & \ddots & \\ \boldsymbol{A}b_{m1} & \dots & \boldsymbol{A}b_{mn}\end{bmatrix}=\boldsymbol{B}\otimes\boldsymbol{A}$

(以下均指右积)

性质 没有交换律, 有结合律, 对加法分配律, 数乘结合律 $\alpha\boldsymbol{A}\otimes\beta\boldsymbol{B}=\alpha\beta(\boldsymbol{A}\otimes\boldsymbol{B})$

转置, 共轭, 厄米, 逆, 广义逆均有分配律 $(\boldsymbol{A}\otimes\boldsymbol{B})^\dagger=\boldsymbol{A}^\dagger\otimes\boldsymbol{B}^\dagger$ (无需换序)

$\mathrm{tr}(\boldsymbol{A}\otimes\boldsymbol{B})=\mathrm{tr}\,\boldsymbol{A}\,\mathrm{tr}\,\boldsymbol{B}$, $\mathrm{rank}(\boldsymbol{A}\otimes\boldsymbol{B})=\mathrm{rank}\,\boldsymbol{A}\,\mathrm{rank}\,\boldsymbol{B}$, $\det(\boldsymbol{A}_{m\times m}\otimes\boldsymbol{B}_{n\times n})=(\det\boldsymbol{A})^n(\det\boldsymbol{B})^m$ (反)

设 $\mathrm{eig}(\boldsymbol{A})=(\lambda_A,\boldsymbol{u}_A)$, $\mathrm{eig}(\boldsymbol{B})=(\lambda_B,\boldsymbol{u}_B)$, 则 $\mathrm{eig}(\boldsymbol{A}\otimes\boldsymbol{B})=(\lambda_A\lambda_B,\boldsymbol{u}_A\otimes\boldsymbol{u}_B)$

例 矢量的外积可用直积表示 $\boldsymbol{a}\boldsymbol{b}^\mathsf{T}=\boldsymbol{a}\otimes\boldsymbol{b}^\mathsf{T}$, $\boldsymbol{a}\otimes\boldsymbol{b}=\mathrm{vec}(\boldsymbol{b}\boldsymbol{a}^\mathsf{T})$

$\boldsymbol{A}\otimes\boldsymbol{O}=\boldsymbol{O}\otimes\boldsymbol{A}=\boldsymbol{O}$, $\boldsymbol{I}_m\otimes\boldsymbol{I}_n=\boldsymbol{I}_{mn}$, $(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D})(\boldsymbol{E}\otimes\boldsymbol{F})=(\boldsymbol{A}\boldsymbol{C}\boldsymbol{E})\otimes(\boldsymbol{B}\boldsymbol{D}\boldsymbol{F})$

$\boldsymbol{K}_{mn}$ 可表示为 $\sum_1^n\boldsymbol{e}_j{}^\mathsf{T}\otimes\boldsymbol{I}_m\otimes\boldsymbol{e}_j$

$\boldsymbol{A}_{m\times n}$, $\boldsymbol{B}_{p\times q}$ 有 $\boldsymbol{K}_{pm}(\boldsymbol{A}\otimes\boldsymbol{B})=(\boldsymbol{B}\otimes\boldsymbol{A})\boldsymbol{K}_{qn}$ $\Leftrightarrow$ $\boldsymbol{K}_{pm}(\boldsymbol{A}\otimes\boldsymbol{B})\boldsymbol{K}_{nq}=(\boldsymbol{B}\otimes\boldsymbol{A})$ $\to$ $\boldsymbol{K}_{pm}(\boldsymbol{A}\otimes\boldsymbol{b})=\boldsymbol{b}\otimes\boldsymbol{A}$

$\boldsymbol{A}_{m\times p}$, $\boldsymbol{B}_{p\times q}$, $\boldsymbol{C}_{q\times n}$ 有 $\mathrm{vec}(\boldsymbol{A}\boldsymbol{B}\boldsymbol{C})=(\boldsymbol{C}^\mathsf{T}\otimes\boldsymbol{A})\,\mathrm{vec}(\boldsymbol{B})=(\boldsymbol{I}_q\otimes\boldsymbol{A}\boldsymbol{B})\,\mathrm{vec}(\boldsymbol{C})=((\boldsymbol{B}\boldsymbol{C})^\mathsf{T}\otimes\boldsymbol{I}_m)\,\mathrm{vec}(\boldsymbol{A})$

定理 厄米/半正定/幺正阵的直积仍厄米/半正定/幺正

元素积

elementwise / Hadamard / Schur product

相同形状矩阵的 元素积 / 哈达玛积 $\boldsymbol{A}_{m\times n}*\boldsymbol{B}_{m\times n}\equiv[a_{ij}b_{ij}]_{m\times n}$

性质 有交换律, 结合律, 加法分配律, 数乘结合律 $\varphi(\boldsymbol{A}*\boldsymbol{B})=(\varphi\boldsymbol{A})*\boldsymbol{B}=\boldsymbol{A}*(\varphi\boldsymbol{B})$

转置, 共轭, 厄米均有分配律

例 $\boldsymbol{A}*\boldsymbol{O}=\boldsymbol{O}*\boldsymbol{A}=\boldsymbol{O}$, $\boldsymbol{A}*\boldsymbol{I}=\boldsymbol{I}*\boldsymbol{A}=\operatorname{diag}(a_{1\sim n})$

$\boldsymbol{A},\boldsymbol{B},\boldsymbol{D}$ 均为 $n\times n$, $\boldsymbol{D}$ 对角, 有 $(\boldsymbol{D}\boldsymbol{A})*(\boldsymbol{B}\boldsymbol{D})=\boldsymbol{D}(\boldsymbol{A}*\boldsymbol{B})\boldsymbol{D}$

$\boldsymbol{A},\boldsymbol{C}$ 为 $m\times m$, $\boldsymbol{B},\boldsymbol{D}$ 为 $n\times n$, 有 $(\boldsymbol{A}\oplus\boldsymbol{B})*(\boldsymbol{C}\oplus\boldsymbol{D})=(\boldsymbol{A}*\boldsymbol{C})\oplus(\boldsymbol{B}*\boldsymbol{D})$

$\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 均为 $m\times n$, 有 $\operatorname{tr}(\boldsymbol{A}^\mathsf{T}(\boldsymbol{B}*\boldsymbol{C}))=\operatorname{tr}((\boldsymbol{A}^\mathsf{T}*\boldsymbol{B}^\mathsf{T})\boldsymbol{C})$

$\boldsymbol{1}_{n\times 1}$, $\boldsymbol{D}=\operatorname{diag}(d_{1\sim m})$, $d_i=\sum a_{ij}$, 有 $\boldsymbol{1}^\mathsf{T}\boldsymbol{A}^\mathsf{T}(\boldsymbol{B}*\boldsymbol{C})\boldsymbol{1}=\operatorname{tr}(\boldsymbol{B}^\mathsf{T}\boldsymbol{D}\boldsymbol{C})$

$\boldsymbol{A},\boldsymbol{B}$ 均为 $n\times n$, $\boldsymbol{D}$ 对角, $\boldsymbol{d}_{n\times 1}=\boldsymbol{D}\boldsymbol{1}$, 有 $\operatorname{tr}(\boldsymbol{A}\boldsymbol{D}\boldsymbol{B}^\mathsf{T}\boldsymbol{D})=\boldsymbol{d}^\mathsf{T}(\boldsymbol{A}*\boldsymbol{B})\boldsymbol{d}$

厄米阵的元素积总仍厄米 $\rightarrow$ 元素积定理 半/正定方阵的元素积仍半/正定 (逆命题不成立)

秩不等式 $\boldsymbol{A},\boldsymbol{B}$ 为方阵, 有 $\operatorname{rank}(\boldsymbol{A}*\boldsymbol{B})\leqslant\operatorname{rank}\boldsymbol{A}\operatorname{rank}\boldsymbol{B}$

本征值不等式 $\boldsymbol{A},\boldsymbol{B}$ 为半正定方阵, μ 是 $\boldsymbol{A}*\boldsymbol{B}$ 的本征值, λ 是 $\boldsymbol{A}\boldsymbol{B}$ 的本征值, 有 $\prod\mu_i\geqslant\prod\lambda_i$

行列式

性质 $|\boldsymbol{A}^\mathsf{T}|=|\boldsymbol{A}|$, $|\boldsymbol{A}^\dagger|=|\boldsymbol{A}|^*$ $\rightarrow$ 厄米阵的行列式为实数

$$\begin{vmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{vmatrix}=\left|\begin{bmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{C} & \boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\end{bmatrix}\begin{bmatrix}\boldsymbol{I} & \boldsymbol{A}^{-1}\boldsymbol{B}\\ \boldsymbol{O} & \boldsymbol{I}\end{bmatrix}\right|\xlongequal{\text{若 }\boldsymbol{A}\text{ 可逆}}|\boldsymbol{A}|\cdot|\boldsymbol{D}-\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}|\xlongequal{\text{若 }\boldsymbol{D}\text{ 可逆}}|\boldsymbol{D}|\cdot|\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}|$$

柯西施瓦茨不等式 $\boldsymbol{A},\boldsymbol{B}$ 均为 $m\times n$ 矩阵, 有 $|\boldsymbol{A}^\dagger\boldsymbol{B}|^2\leqslant|\boldsymbol{A}^\dagger\boldsymbol{A}|\cdot|\boldsymbol{B}^\dagger\boldsymbol{B}|$

Hadamard

哈达玛不等式 (1893) $|\boldsymbol{A}|\leqslant\prod_i^n\|\boldsymbol{a}_j\|$, 列矢量 $\|\boldsymbol{a}_j\|=(\sum_j^n|a_{ij}|^2)^{1/2}$ 「平行六面体各边正交时体积最大」

「$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{L}^\dagger$ 」若 $\boldsymbol{A}$ 半正定, 有 $|\boldsymbol{A}|=|\boldsymbol{L}|^2\leqslant\prod\|\boldsymbol{l}_j\|^2=\prod a_{ii}$, 若正定则仅当 $\boldsymbol{A}$ 对角时取等号

Fischer

费舍尔不等式 分块阵正定, $\boldsymbol{A},\boldsymbol{C}\neq\boldsymbol{O}$, 则 $\begin{vmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{B}^\dagger & \boldsymbol{C}\end{vmatrix}=|\boldsymbol{A}|\cdot|\boldsymbol{C}-\boldsymbol{B}^\dagger\boldsymbol{A}^{-1}\boldsymbol{B}|\leqslant|\boldsymbol{A}|\cdot|\boldsymbol{C}|$

Minkowski

闵可夫斯基不等式 $\boldsymbol{A},\boldsymbol{B}\neq\boldsymbol{O}$ 为 n 阶半正定方阵, 则 $\sqrt[n]{|\boldsymbol{A}+\boldsymbol{B}|}\geqslant\sqrt[n]{|\boldsymbol{A}|}+\sqrt[n]{|\boldsymbol{B}|}$, 仅当 $\boldsymbol{B}=\varphi\boldsymbol{A}$ 取等

「齐次两边可以约, 可取 $\boldsymbol{A}=\boldsymbol{I}$, $\lambda_i^{(\boldsymbol{B})}>0$, 证 $\prod(1+\lambda_i)\geqslant(1+\sqrt[n]{\prod\lambda_i})^n$ 初等」

Oppenheim

奥本海姆不等式 $\boldsymbol{A},\boldsymbol{B}$ 为半正定方阵, 有 $\prod a_{ii}\prod b_{ii}\geqslant|\boldsymbol{A}*\boldsymbol{B}|\geqslant(\prod a_{ii})|\boldsymbol{B}|\geqslant|\boldsymbol{A}|\cdot|\boldsymbol{B}|$ 「对 n 用归纳法」

Ostrowski-Taussky

奥斯特洛夫斯基定理 若 $\boldsymbol{H}\equiv\frac{1}{2}(\boldsymbol{A}+\boldsymbol{A}^\dagger)$ 正定, 则 $|\boldsymbol{H}|\leqslant|\boldsymbol{A}|$, 仅当 $\boldsymbol{A}$ 为厄米阵时取等号

秩

n 维行/列矢量组, 最多存在 n 个线性无关的矢量 $\rightarrow\operatorname{rank}\boldsymbol{A}\leqslant\min(m,n)$, 取等号则称 满秩 〈线代〉

性质 $\operatorname{rank}\boldsymbol{A}=\operatorname{rank}\boldsymbol{A}^\mathsf{T}=\operatorname{rank}\boldsymbol{A}^*=\operatorname{rank}\boldsymbol{A}^\dagger$, 若 $\varphi\neq 0$ 则 $\operatorname{rank}(\varphi\boldsymbol{A})=\operatorname{rank}\boldsymbol{A}$,

子矩阵的秩 $\leqslant$ 原矩阵的秩, $\boldsymbol{A}$ 左乘列满秩 $\boldsymbol{P}_{m\times m}$ 或右乘行满秩 $\boldsymbol{Q}_{n\times n}$ 秩不变

equivalence

$\rightarrow\operatorname{rank}\boldsymbol{A}_{m\times n}=\operatorname{rank}\boldsymbol{B}_{m\times n}\Leftrightarrow\exists$ 可逆矩阵 $\boldsymbol{P}_{m\times m},\boldsymbol{Q}_{n\times n}$ 使 $\boldsymbol{B}=\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q}\Leftrightarrow$ 两矩阵 等价 $\boldsymbol{A}\cong\boldsymbol{B}$

$\operatorname{rank}\boldsymbol{A}=r\Leftrightarrow$ 存在 $\boldsymbol{A}$ 的一个 $r\times r$ 子矩阵的行列式非零, 且所有 $(r+1)\times(r+1)$ 子矩阵的行列式为零

full-rank factorization

满秩分解 $\operatorname{rank}\boldsymbol{A}_{m\times n}=k\Leftrightarrow$ 存在列满秩的 $\boldsymbol{P}_{m\times k}$, $\boldsymbol{Q}_{n\times k}$, 使 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{Q}^\mathsf{T}\rightarrow$ 秩 1 阵总可表为 $\boldsymbol{x}\boldsymbol{y}^\mathsf{T}$

(由迷向矢量构造的秩 1 对称阵 $\boldsymbol{v}\boldsymbol{v}^\mathsf{T}$ 不能对角化)

rank-sum inequality

和秩不等式 $|\operatorname{rank}\boldsymbol{A}-\operatorname{rank}\boldsymbol{B}|\leqslant\operatorname{rank}(\boldsymbol{A}+\boldsymbol{B})\leqslant\operatorname{rank}[\boldsymbol{A}\ \boldsymbol{B}]\leqslant\operatorname{rank}\boldsymbol{A}+\operatorname{rank}\boldsymbol{B}$

「设 $\boldsymbol{A},\boldsymbol{B}$ 的极大无关组为 $(\boldsymbol{a}_{1\sim r_a}),(\boldsymbol{b}_{1\sim r_b})$, 则仅当它们无关时右端不等式取等号」

Sylvester inequality

「若 $\operatorname{rank}\boldsymbol{B}=1$ 」$\rightarrow$ 改变矩阵的一行/列最多只能使秩 ± 1

西尔维斯特不等式 $\operatorname{rank}\boldsymbol{A}+\operatorname{rank}\boldsymbol{B}-n\leqslant\operatorname{rank}(\boldsymbol{A}\boldsymbol{B})\leqslant\min(\operatorname{rank}\boldsymbol{A},\operatorname{rank}\boldsymbol{B})$

「左边: 构造 $\boldsymbol{C}=\begin{bmatrix}\boldsymbol{I}_n & \boldsymbol{B}\\ \boldsymbol{A} & \boldsymbol{O}\end{bmatrix}$, 则 $\geqslant\operatorname{rank}\boldsymbol{A}+\operatorname{rank}\boldsymbol{B}$, 可初等变换为 $\begin{bmatrix}\boldsymbol{I}_n & \boldsymbol{O}\\ -\boldsymbol{A} & \boldsymbol{I}_m\end{bmatrix}\boldsymbol{C}\begin{bmatrix}\boldsymbol{I}_n & -\boldsymbol{B}\\ \boldsymbol{O} & \boldsymbol{I}_s\end{bmatrix}=\begin{bmatrix}\boldsymbol{I}_n & \boldsymbol{O}\\ \boldsymbol{O} & -\boldsymbol{A}\boldsymbol{B}\end{bmatrix}$ 」

「右边: $\boldsymbol{A}\boldsymbol{B}$ 的列可由 $\boldsymbol{A}$ 的列线性表出, 故 $\operatorname{rank}(\boldsymbol{A}\boldsymbol{B})\leqslant\operatorname{rank}\boldsymbol{A}$, $\boldsymbol{B}$ 同理」

例 $\boldsymbol{A}$ 为高矩阵, 则 $\boldsymbol{A}\boldsymbol{A}^\top$ 必不满秩 定理 $\operatorname{rank}\boldsymbol{A}\geqslant\operatorname{rank}(\boldsymbol{A}^\top\boldsymbol{A})$

推论 $\operatorname{rank}\boldsymbol{A}=n-1$ 时 $\operatorname{rank}\boldsymbol{A}^\#=1$, $\operatorname{rank}\boldsymbol{A}<n-1$ 时 $\operatorname{rank}\boldsymbol{A}^\#=0$

Frobenius inequality
弗罗贝尼乌斯不等式 $\operatorname{rank}(\boldsymbol{AB})+\operatorname{rank}(\boldsymbol{BC})\leqslant\operatorname{rank}\boldsymbol{B}+\operatorname{rank}(\boldsymbol{ABC})$ 仅当 $\boldsymbol{B}=\boldsymbol{XAB}+\boldsymbol{BCY}$ 时取等

迹

方阵的对角元之和称为 迹 (trace) $\operatorname{tr}\boldsymbol{A}=\sum a_{ii}=\sum\lambda_i$

性质 齐次叠加性 $\operatorname{tr}(\varphi_1\boldsymbol{A}+\varphi_2\boldsymbol{B})=\varphi_1\operatorname{tr}\boldsymbol{A}+\varphi_2\operatorname{tr}\boldsymbol{B}$, $\operatorname{tr}\boldsymbol{A}^\top=\operatorname{tr}\boldsymbol{A}$, $\operatorname{tr}\boldsymbol{A}^\dagger=\operatorname{tr}\boldsymbol{A}^*=(\operatorname{tr}\boldsymbol{A})^*$

定理 $\boldsymbol{A}_{m\times n}$, $\boldsymbol{B}_{n\times m}$ 有 $\operatorname{tr}(\boldsymbol{AB})=\operatorname{tr}(\boldsymbol{BA})\rightarrow\operatorname{tr}(\boldsymbol{ABC})=\operatorname{tr}(\boldsymbol{BCA})=\operatorname{tr}(\boldsymbol{CAB})$ (没有 $=\operatorname{tr}(\boldsymbol{ACB})$)

$\rightarrow$ 若 $\boldsymbol{B}$ 可逆, 有 $\operatorname{tr}\boldsymbol{A}=\operatorname{tr}(\boldsymbol{ABB}^{-1})=\operatorname{tr}(\boldsymbol{B}^{-1}\boldsymbol{AB})=\operatorname{tr}(\boldsymbol{BAB}^{-1})$

例 $\boldsymbol{x}^\dagger\boldsymbol{A}\boldsymbol{x}=\operatorname{tr}(\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^\dagger)$, $\boldsymbol{y}^\dagger\boldsymbol{x}=\operatorname{tr}(\boldsymbol{x}\boldsymbol{y}^\dagger)$

$\operatorname{tr}(\boldsymbol{A}^\top\boldsymbol{B})=\operatorname{vec}(\boldsymbol{A})^\top\operatorname{vec}(\boldsymbol{B})=\sum_{\text{所有元素}}\operatorname{vec}(\boldsymbol{A}*\boldsymbol{B})$, $\operatorname{tr}(\boldsymbol{ABCD})=\operatorname{vec}(\boldsymbol{D})^\top(\boldsymbol{A}\otimes\boldsymbol{C}^\top)\operatorname{vec}(\boldsymbol{B}^\top)$

柯西施瓦茨不等式 $\boldsymbol{A}_{m\times n}$, $\boldsymbol{B}_{m\times n}$ 有 $\operatorname{tr}((\boldsymbol{A}^\top\boldsymbol{B})^2)\leqslant\operatorname{tr}(\boldsymbol{A}^\top\boldsymbol{A})\operatorname{tr}(\boldsymbol{B}^\top\boldsymbol{B})$, 当且仅当 $\boldsymbol{A}=\varphi\boldsymbol{B}$ 时取等号

$\operatorname{tr}((\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{A}+\boldsymbol{B})^\top)\leqslant2[\operatorname{tr}(\boldsymbol{AA}^\top)+\operatorname{tr}(\boldsymbol{BB}^\top)]$, $\boldsymbol{A},\boldsymbol{B}$ 为同阶对称方阵, 有 $\operatorname{tr}(\boldsymbol{AB})\leqslant\frac{1}{2}\operatorname{tr}(\boldsymbol{A}^2+\boldsymbol{B}^2)$

范数

① 正定性 ② 齐次性 $\|\varphi\boldsymbol{x}\|=|\varphi|\,\|\boldsymbol{x}\|$ ③ 三角不等式 $\|\boldsymbol{x}+\boldsymbol{y}\|\leqslant\|\boldsymbol{x}\|+\|\boldsymbol{y}\|$, 将一个矢量映射到一个实数, 满足 ① ② ③ 构成矢量 范数 (norm) , 去掉定性称为 半范数 / 伪范数 (seminorm / pseudo-norm) (处理迷向矢量) 减弱 ③ 称为 准范数 (pre-norm)

推论 半/范数的正数倍, 两种半/范数之和仍是半/范数, 做可逆线性变换 $\|\boldsymbol{Ax}\|$ 可构造新范数

Hölder
l_p 范数 $\|\boldsymbol{x}\|_p=(\sum|x_i|^p)^{1/p}$, $C[a,b]$ 上 $\|f\|_p\equiv\left[\int_a^b|f(t)|^p\mathrm{d}t\right]^{1/p}$, $p\geqslant1$ 是范数, $0<p<1$ 是拟范数

(三角不等式就是 Minkowski)

Euclidean
l_2 / 欧氏范数 $\|\boldsymbol{x}\|_2=\sqrt{\sum x_i^2}$ 由欧氏内积诱导 (除原点外均可微, 常用于优化问题)

l_0 范数 $\|\boldsymbol{x}\|_0\equiv$ 非零元素个数 (用于稀疏表示) l_1 / 和范数 (sum / Manhattan / taxicab) $\|\boldsymbol{x}\|_1\equiv\sum|x_i|$ (小范围可微, 统计学常用)

max norm
l_∞ 范数 $\|\boldsymbol{x}\|_\infty\equiv\max(|x_i|)$ (直接检验收敛性, 实际用的不方便)

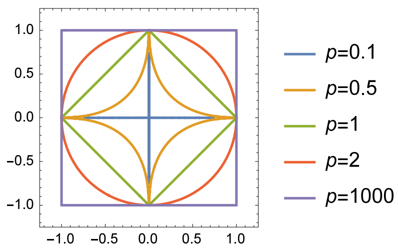
若 $\|x\|=\sqrt{\langle\boldsymbol{x},\boldsymbol{x}\rangle}$ 则称为由内积 诱导 出的范数 例 l_1 , l_∞ 都不能由内积诱导

由内积诱导的范数 $\Leftrightarrow$ 平行四边形法则 $\Leftrightarrow$ 极化恒等式 〈 线代 〉

定理 对于半范数有 $|\,\|x\|-\|y\|\,|\leqslant\|\boldsymbol{x}-\boldsymbol{y}\|$, $\|\boldsymbol{x}\|\leqslant\sum|x_i|\,\|\boldsymbol{e}_i\|$ 例 $\|\boldsymbol{x}\|\equiv\sum x_i$ 是半范数

若 $\forall\boldsymbol{x}\in V$ 有 $\|\boldsymbol{Ax}\|=\|\boldsymbol{x}\|$, 则称方阵 $\boldsymbol{A}$ 是关于该范数的 等距变换 性质 等距变换必可逆

例 l_2 范数酉不变, l_1 范数的等距变换是由模 1 复数构成的置换阵



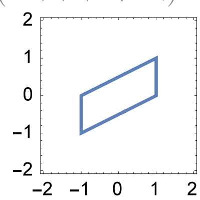
best bounds
$1\leqslant p<q<\infty$, 则 最佳界 $\|\boldsymbol{x}\|_q\leqslant\|\boldsymbol{x}\|_p\leqslant n^{(\frac{1}{p}-\frac{1}{q})}\|\boldsymbol{x}\|_1$

$B_p=\{\|\boldsymbol{v}\|\leqslant1\}$ 称为 单位球 定理 $\forall\boldsymbol{x}\in V$ 有 $\|\boldsymbol{x}\|_q\leqslant\|\boldsymbol{x}\|_p\Leftrightarrow B_p\subset B_q$

若 $\forall\boldsymbol{x}\in B$, $|w|=1$ 有 $w\boldsymbol{x}\in B$ 则称 B 是 均衡的

有限非零维中, 集合 B 是范数的单位球 $\Leftrightarrow$ 紧集, 以 $\mathbf{0}$ 为内点(正定) , 均衡集(齐次) , 凸集(三角不等式)

记 $|\boldsymbol{x}|$ 为每个元取绝对值, $|\boldsymbol{x}|<|\boldsymbol{y}|$ 指每个元都小于, 若 $\|\boldsymbol{x}\|=\|\,|\boldsymbol{x}|\,\|$ 称 绝对



weakly monotone
$|\boldsymbol{x}|\leqslant|\boldsymbol{y}|\Rightarrow\|\boldsymbol{x}\|\leqslant\|\boldsymbol{y}\|$ 称 单调 , $\Rightarrow\|[x_i,0,x_j]\|\leqslant\|[x_i,x_k,x_j]\|$ 称 弱单调

(将球面上一点的一个坐标变到零, 整个线段在单位球内) 定理 对有限维, 单调 $\Leftrightarrow$ 绝对

例 对二维矢量 $|x_1-x_2|+|x_2|$ 是范数, 不弱单调

若序列对范数 $\|\boldsymbol{x}_i\|_\alpha$ 收敛于 $\boldsymbol{x}$ 就也对范数 $\|\boldsymbol{x}_i\|_\beta$ 收敛于 $\boldsymbol{x}$, 则称两种范数 等价

定理 有限维实或复线性空间, 所有范数都等价 (无穷维不一定, 如 $\|\delta\|_\infty=\infty$)

$\rightarrow$ 任意范数或准范数的单位球及单位球面是紧集

矩阵范数

generalized matrix norm　submultiplicativity　matrix / ring norm

满足正定齐次三角不等式称为 广义矩阵范数，还满足 次乘性 $\|\boldsymbol{AB}\|\leqslant\|\boldsymbol{A}\|\,\|\boldsymbol{B}\|$ 才称 矩阵范数

推论 $\|\boldsymbol{I}\|\geqslant 1$，$\|\boldsymbol{A}^k\|\leqslant\|\boldsymbol{A}\|^k$ $(k\in\mathbb{N}_+)$

Frobenius / Schur / Hilbert-Schmidt

例 l_2 范数 可沿用到矩阵上 $\|\boldsymbol{A}\|_\mathrm{F}=\sqrt{\sum_{ij}|a_{ij}|^2}=\sqrt{\mathrm{tr}(\boldsymbol{A}^\dagger\boldsymbol{A})}=\sqrt{\mathrm{tr}(\boldsymbol{A}\boldsymbol{A}^\dagger)}=\sqrt{\sum\sigma_i^2(\boldsymbol{A})}$ **性质** 绝对，酉不变，$\|\boldsymbol{AB}\|_\mathrm{F}=\|\boldsymbol{A}\|_\mathrm{F}\,\|\boldsymbol{B}\|_\mathrm{F}$ 「次乘性就是柯西施瓦茨 $\sum_{ij}|\sum_k a_{ik}b_{kj}|^2\leqslant\sum_{ij}(\sum_k|a_{ik}|^2)(\sum_l|b_{lj}|^2)=(\sum_{ik}|a_{ik}|^2)(\sum_{lj}|b_{lj}|^2)$ 」l_1 矢量范数可沿用「证次乘用三角不等式 + 缩放 $\|AB\|_1=\sum_{ij}|\sum_k a_{ik}b_{kj}|\leqslant\sum_{ijk}|a_{ik}b_{kj}|\leqslant\sum_{ijkl}|a_{ik}b_{lj}|=(\sum_{ik}|a_{ik}|)(\sum_{lj}|b_{lj}|)$ 」l_∞ 不满足次乘性，需做修改 $\|\boldsymbol{A}\|_\infty\equiv n(\max|a_{ij}|)$

induced

$\|\boldsymbol{A}\|=\max\limits_{\|\boldsymbol{x}\|=1}\|\boldsymbol{Ax}\|$ 称为由矢量范数 诱导 的矩阵范数 $\Rightarrow\|\boldsymbol{I}\|=1$

maximum row sum

l_1 诱导 最大列和范数 $\|\boldsymbol{A}\|_1=\max\limits_{1\leqslant j\leqslant n}\sum\limits_{i=1}^n|a_{ij}|$，$l_\infty$ 诱导 最大行和范数 $\|\boldsymbol{A}\|_\infty=\max\limits_{1\leqslant i\leqslant n}\sum\limits_{j=1}^n|a_{ij}|$

spectral norm

l_2 诱导 谱范数 $\|\boldsymbol{A}\|_2=\max\sqrt{\lambda^{(\boldsymbol{A}^\dagger\boldsymbol{A})}}=\sqrt{\rho(\boldsymbol{A}^\dagger\boldsymbol{A})}=\sigma_1(\boldsymbol{A})$ **性质** 酉不变，厄米 $\|\boldsymbol{A}^\dagger\|=\|\boldsymbol{A}\|$

谱半径 $\rho(\boldsymbol{A})\equiv\max|\lambda_i|$ 为非负实数 **性质** 连续函数，有正性，齐次性，没有定性，叠加性，次乘性

定理 $\boldsymbol{A}$ 为方阵，任意矩阵范数 $\rho(\boldsymbol{A})\leqslant\|\boldsymbol{A}\|$「$|\lambda|\,\|\boldsymbol{U}\|=\|\lambda\boldsymbol{U}\|=\|\boldsymbol{AU}\|\leqslant\|\boldsymbol{A}\|\,\|\boldsymbol{U}\|$」

矩阵 收敛 $\lim\limits_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{O}\Leftrightarrow\rho(\boldsymbol{A})<1$ （谱半径本身不是范数，但给出最大下界）

$\|\boldsymbol{A}\|_2^2=\rho(\boldsymbol{A}^\dagger\boldsymbol{A})\leqslant\left\|\boldsymbol{A}^\dagger\boldsymbol{A}\right\|_1\leqslant\|\boldsymbol{A}\|_\infty\|\boldsymbol{A}\|_1$，右边可推广为 对偶范数 之积

推论 $\rho(\boldsymbol{A})=\lim\limits_{k\to\infty}\left\|\boldsymbol{A}^k\right\|^{1/k}\to\mathrm{e}^{\boldsymbol{A}}=\sum_{k=0}^\infty\frac{1}{k!}\boldsymbol{A}^k$ 对所有方阵都有定义

例 对 $|z|<1$ 有 $(1-z)^{-1}=1+z+z^2+\ldots$，对 $\|\boldsymbol{A}\|<1$ 有 $(\boldsymbol{I}-\boldsymbol{A})^{-1}=\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2+\ldots$ （任何范数均如此）

方阵是 严格对角占优阵 $|a_{ii}|>\sum_j|a_{ij}|$ $(j=1\sim n,j\neq i)$ $\Rightarrow$ 可逆 (Levy-Desplanques)

本征值

定理 对于任意实矩阵，若有复本征值则必成对共轭出现「$\boldsymbol{Ax}=\lambda\boldsymbol{x}$ 两边取共轭 $\boldsymbol{Ax}^*=\lambda^*\boldsymbol{x}^*$」

$\boldsymbol{A}_{m\times n}$，$\boldsymbol{B}_{n\times m}$，$n\geqslant m$，

若 $\mathrm{eig}(\boldsymbol{AB})=\lambda_{1\sim m}$，则 $\mathrm{eig}(\boldsymbol{BA})=\lambda_{1\sim m}$ 和 $(n-m)$ 个 0　$\begin{bmatrix}\boldsymbol{I}_m&-\boldsymbol{A}\\\boldsymbol{O}&\boldsymbol{I}_n\end{bmatrix}\begin{bmatrix}\boldsymbol{AB}&\boldsymbol{O}\\\boldsymbol{B}&\boldsymbol{O}_n\end{bmatrix}\begin{bmatrix}\boldsymbol{I}_m&\boldsymbol{A}\\\boldsymbol{O}&\boldsymbol{I}_n\end{bmatrix}=\begin{bmatrix}\boldsymbol{O}_m&\boldsymbol{O}\\\boldsymbol{B}&\boldsymbol{BA}\end{bmatrix}$

$\to$ $\boldsymbol{A},\boldsymbol{B}$ 为方阵，至少有一者可逆，则 $\boldsymbol{AB}\sim\boldsymbol{BA}$ （本征值相同，本征矢不同）

Cauchy's determinant identity

$\to$ 柯西行列式 $|\boldsymbol{A}+\boldsymbol{xy}^\mathsf{T}|=|\boldsymbol{A}||\boldsymbol{I}+\boldsymbol{A}^{-1}\boldsymbol{xy}^\mathsf{T}|=|\boldsymbol{A}|\prod\lambda_i^{(\boldsymbol{I}+\boldsymbol{A}^{-1}\boldsymbol{xy}^\mathsf{T})}=|\boldsymbol{A}|\prod(1+\lambda_i^{(\boldsymbol{A}^{-1}\boldsymbol{xy}^\mathsf{T})})$

$=|\boldsymbol{A}|(1+\boldsymbol{y}^\mathsf{T}\boldsymbol{A}^{-1}\boldsymbol{x})=|\boldsymbol{A}|+\boldsymbol{x}^\mathsf{T}\boldsymbol{A}^{\#}\boldsymbol{y}$

$\to$ 求常对角或汉克尔阵的本征值，可分解为矢量然后换序计算　**例** $\boldsymbol{A}=\left[(i-j)\right]_{n\times n}=\begin{bmatrix}0&-1&-2&\ldots\\1&0&-1&\ldots\\2&1&0&\ddots\end{bmatrix}$

$=\left[\boldsymbol{v}\ \ -\boldsymbol{1}_n\right]\left[\boldsymbol{1}_n\ \ \boldsymbol{v}\right]^\mathsf{T}$，$\boldsymbol{v}=\left[1\ 2\ \ldots\ n\right]^\mathsf{T}$，换序相乘求得 $\lambda=\pm\frac{n\mathrm{i}}{2}\sqrt{\frac{n^2-1}{3}}$ 和 $n-2$ 个 0

瑞利商

Rayleigh-Ritz　Rayleigh quotient

瑞利里兹定理 $\boldsymbol{A}$ 厄米，设本征值按 $\lambda_1\leqslant\ldots\leqslant\lambda_n$ 排序，则 $\forall\boldsymbol{x}$ 有 $\lambda_1\leqslant R\leqslant\lambda_n$，瑞利商 $R(\boldsymbol{x},\boldsymbol{A})\equiv\dfrac{\boldsymbol{x}^\dagger\boldsymbol{Ax}}{\boldsymbol{x}^\dagger\boldsymbol{x}}$

「设 $\boldsymbol{A}=\boldsymbol{U}^\dagger\boldsymbol{\Lambda U}$，有 $\boldsymbol{x}^\dagger\boldsymbol{Ax}=(\boldsymbol{Ux})^\dagger\boldsymbol{\Lambda}(\boldsymbol{Ux})=\sum\lambda_i\left\|(\boldsymbol{Ux})_i\right\|^2$ 」

性质 齐次性 $R(\alpha\boldsymbol{x},\beta\boldsymbol{A})=\beta R(\boldsymbol{x},\boldsymbol{A})$ 平移不变性 $R(\boldsymbol{x},\boldsymbol{A}-\alpha\boldsymbol{I})=\beta R(\boldsymbol{x},\boldsymbol{A})-\alpha$ 正交性 $\boldsymbol{x}\perp(\boldsymbol{A}-R\boldsymbol{I})\boldsymbol{x}$

有界性 当 $\boldsymbol{x}$ 取遍所有非零矢量，R 为复平面上的一个闭合有界凸区域，称为 $\boldsymbol{A}$ 的 值域，若 $\boldsymbol{A}$ 厄米则区域是个闭区间 $[\lambda_1,\lambda_n]$ （除了本征方程，厄米阵的本征值还可代表一系列条件优化问题的解〈线代〉）

Weyl

外尔定理 $\boldsymbol{A},\boldsymbol{B}$ 厄米，本征值 $\lambda^{(\boldsymbol{A})},\lambda^{(\boldsymbol{B})},\lambda^{(\boldsymbol{A}+\boldsymbol{B})}$ 均按递增排列，则 $\lambda_{1+s}^{(\boldsymbol{A})}+\lambda_{i-s}^{(\boldsymbol{B})}\leqslant\lambda_i^{(\boldsymbol{A}+\boldsymbol{B})}\leqslant\lambda_{i+t}^{(\boldsymbol{A})}+\lambda_{n-t}^{(\boldsymbol{B})}$，其中 $s=0\sim i,t=0\sim(n-i)$ 任意整数 $\to$ 本征值扰动定理 $|\lambda_k^{(\boldsymbol{A}+\boldsymbol{B})}-\lambda_k^{(\boldsymbol{A})}|\leqslant\rho(\boldsymbol{B})$

单调性定理 厄米阵加上正定阵后，所有本征值比原来大 交错本征值定理 给厄米阵加上秩 1 厄米阵，或加边得 $(n+1)\times(n+1)$ 厄米阵，则新旧矩阵的本征值必交错 $\lambda_1'\leqslant\lambda_1\leqslant\ldots\leqslant\lambda_n'\leqslant\lambda_n\leqslant\lambda_{n+1}'$

包含原理 记 $\boldsymbol{A}_r$ 为厄米阵 $\boldsymbol{A}$ 的 $r\times r$ 顺序主子式，有 $\lambda_k^{(\boldsymbol{A})}\leqslant\lambda_k^{(\boldsymbol{A}_r)}\leqslant\lambda_{k+n-r}^{(\boldsymbol{A})}$，其中 $k=1\sim r$ 整数

$\to$ 庞加莱分离定理 上式 $\boldsymbol{A}_r$ 换成 $\boldsymbol{B}_r\equiv\left[\boldsymbol{u}_i^\dagger\boldsymbol{Au}_j\right]$，$(\boldsymbol{u}_{1\sim r})$ 正交归一

奇异值

「正交对角化只适用于对称阵〈线代〉对任意形状的矩阵，可用两个正交阵来对角化」 $A\begin{bmatrix}v_1 & \dots\end{bmatrix}=\begin{bmatrix}u_1 & \dots\end{bmatrix}\begin{bmatrix}\sigma_1 & \\ & \ddots\end{bmatrix}$

Singular Value Decomposition
奇异值分解 (Beltrami 1873) $A_{m\times n}=U_{m\times m}\Sigma_{m\times n}V_{n\times n}^{\top}$

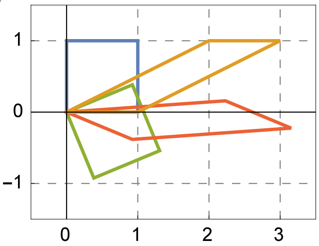
→ $Av_i=\sigma_i u_i$ （A 把列空间正交基 (v_i) 变换成另一组正交基 (u_i)，伸缩因子 σ_i）

「$AA^{\top}=U\Sigma\Sigma^{\top}U^{\top}$」 $AA^{\top}$ 的本征矢就是 **左奇异矢** u_i

「$A^{\top}A=V\Sigma^{\top}\Sigma V^{\top}$」 $A^{\top}A$ 的本征矢就是 **右奇异矢** v_i

$AA^{\top}$ 和 $A^{\top}A$ 本征值相同，算术平方根就是奇异值，习惯从大到小排列

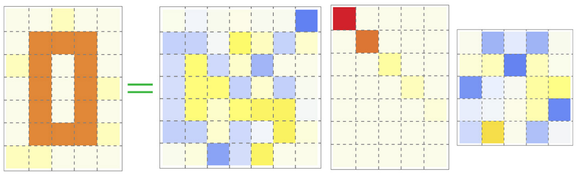
例 $\begin{bmatrix}1&2\\0&1\end{bmatrix}=\begin{bmatrix}\frac{\sqrt{2+\sqrt2}}{2}&\frac{\sqrt{2-\sqrt2}}{-2}\\ \frac{\sqrt{2-\sqrt2}}{2}&\frac{\sqrt{2+\sqrt2}}{2}\end{bmatrix}\begin{bmatrix}\sqrt{3+2\sqrt2}&0\\0&\sqrt{3-2\sqrt2}\end{bmatrix}\begin{bmatrix}\frac{\sqrt{2-\sqrt2}}{2}&\frac{\sqrt{2+\sqrt2}}{2}\\ \frac{\sqrt{2+\sqrt2}}{-2}&\frac{\sqrt{2-\sqrt2}}{2}\end{bmatrix}$



蓝色 → 转 $\frac{3}{8}\pi$ → 绿色 → 纯缩放 → 红色 → 转 $-\frac{\pi}{8}$ → 橙色

SVD 还可写成 $A=u_1\sigma_1v_1^{\top}+u_2\sigma_2v_2^{\top}+\dots$ → $u_i^{\top}Av_i=\sigma_i$

compression noise reduction
截取最大的几个奇异值可实现数据的 **压缩** 和 **降噪**



Principle Component Analysis
主成份分析 估计高维数据的低维子空间，PCA 也使用了 SVD 去检测数据间依赖和冗余信息

性质 A 和 $A^{\dagger}$ 的奇异值相同，$|A|=\prod\sigma_i$，$\mathrm{tr}(A^{\dagger}A)=\sum\sigma_i^2$，$\mathrm{rank}\,A=$ 非零奇异值的个数

对于方阵，设本征值 $\lambda_{1\sim n}$ 按 $|\lambda_1|\geqslant\dots\geqslant|\lambda_n|$ 排列，特征值 $\sigma_{1\sim n}$ 按 $\sigma_1\geqslant\dots\geqslant\sigma_n\geqslant0$，则 $\sigma_1\geqslant|\lambda_i|\geqslant\sigma_n$

按 $\sigma_1\geqslant\dots\geqslant\sigma_{\min(m,n)}$ 排列，则 $\sigma_{i+j-1}(A+B)\leqslant\sigma_i(A)+\sigma_j(B)$，$\sigma_{i+j-1}(AB^{\dagger})\leqslant\sigma_i(A)\sigma_j(B)$

设 B 是 A 删去一行/列得到的矩阵，则 $\sigma_1(A)\geqslant\sigma_1(B)\geqslant\sigma_2(A)\geqslant\sigma_2(B)\geqslant\dots\geqslant0$

polar decomposition
「类似 $z=re^{i\theta}$」 **极分解** 宽矩阵 $A=PU$，$P=\sqrt{AA^{\dagger}}$，高矩阵 $A=UQ$，$Q=\sqrt{A^{\dagger}A}$，P,Q 半正定，

「$(U\Sigma U^{\dagger})(UV^{\dagger})$」 若 A 行/列满秩则幺正阵 U 唯一，方阵 $A=PU=UQ$，A 为实则 P,Q,U 均为实

条件数

condition number
A 作系数矩阵时，**条件数** $\mathrm{cond}(A)$ 或 $\kappa(A)\equiv\|A\|\,\|A^{-1}\|$ 「设真值是 $Ax=b$，若 b 有扰动，解得 $A(x+\delta x)=b+\delta b$ → $\delta x=A^{-1}\delta b$，由矩阵范数次乘性 $\|\delta x\|\leqslant\|A^{-1}\|\,\|\delta b\|$，$\|A\|\,\|x\|\geqslant\|b\|$，相除得 $\frac{\|\delta x\|}{\|x\|}\leqslant\kappa\frac{\|\delta b\|}{\|b\|}$，类似地若 A 有扰动，由 $(A+\delta A)(x+\delta x)=b$ 可得 $\frac{\|\delta x\|}{\|x+\delta x\|}\leqslant\kappa\frac{\|\delta A\|}{\|A\|}$」 表征误差传播

ill conditioned well perfectly
越大则 A 越接近奇异，称为 **病态**，接近 1 称为 **良态**，$=1$ 称为 **优态** **例** $\kappa(I)=1$，$\kappa(\text{奇异阵})=\infty$

对于超定方程「$A^{\dagger}A=V\Sigma^2V^{\dagger}$」 $\kappa(A^{\dagger}A)=\kappa(A)^2=\sigma_1^2/\sigma_n^2$ → $\kappa(A)=\sigma_1/\sigma_{\min(m,n)}\geqslant|\lambda_1|/|\lambda_n|\geqslant1$

任意矩阵的单个本征值的 **条件数** $\mathrm{cond}(\lambda)\equiv\frac{1}{\cos\theta}$，$\theta$ 为左右本征矢夹角，表征本征值受阵元扰动倍数

伪逆

left inverse right inverse
对于宽矩阵可定义 **左逆** $LA=I$ （仅 $m\geqslant n$ 才可能有左逆），高矩阵可定义 **右逆** $AR=I$

left pseudo inverse right pseudo inverse
若 A 列满秩，则 $L=(A^{\dagger}A)^{-1}A^{\dagger}$ 唯一，称为 **左伪逆**，$R=A^{\dagger}(AA^{\dagger})^{-1}$ 称为 **右伪逆**

AL 是往列空间投影，即求最小二乘解，RA 是往行空间投影

(零空间导致不可逆，若不管零空间，A 是完美的行空间到列空间的双射（维数相同），此即伪逆)

well-determined
适定 n 维方阵 $\mathrm{rank}\,A=n$ ⇔ A 可逆 → 有唯一解 $x=A^{-1}b$

under-determined
欠定 独立方程个数 < 独立未知参数个数，有无穷多组解 → **最小范数解** $x=Rb$

over-determined
超定 独立方程个数 > 独立未知参数个数，可能无解，→ **最小二乘解** $x=Lb$

minimum norm least squares solution
若秩亏缺，则最小二乘解不唯一，其中 **最小范数最小二乘解** 是 $A^{-}b$

广义逆

generalized inverse

秩亏缺矩阵 $\boldsymbol{A}_{m\times n}$ 的逆称为 广义逆矩阵 $\boldsymbol{A}^-_{n\times m}$ (Moore-Penrose 条件, 分别 1935,1955)

「两个矩阵之积不可能 $=\boldsymbol{I}$，故需三个矩阵, 考虑矩阵方程 $\boldsymbol{Ax}=\boldsymbol{y}$，两边左乘 $\boldsymbol{AA}^-$，将 $\boldsymbol{x}=\boldsymbol{A}^-\boldsymbol{y}$ 代入」

① $\boldsymbol{AA}^-\boldsymbol{A}=\boldsymbol{A}$「这样的广义逆 $\boldsymbol{A}_{n\times m}$ 不唯一, 要保证 $\boldsymbol{A}$ 也是 $\boldsymbol{A}^-$ 的逆, 再考虑方程 $\boldsymbol{A}^-\boldsymbol{y}=\boldsymbol{x}$ 即可」

② $\boldsymbol{A}^-\boldsymbol{AA}^-=\boldsymbol{A}^-$「然后还要兼容左/右伪逆, $\boldsymbol{AL}=\boldsymbol{A}(\boldsymbol{A}^\dagger\boldsymbol{A})^{-1}\boldsymbol{A}^\dagger=(\boldsymbol{AL})^\dagger$」③ $\boldsymbol{AA}^-$ 和 ④ $\boldsymbol{A}^-\boldsymbol{A}$ 厄米逆, 左/右伪逆均满足全部 4 个条件, 称为 MP 广义逆 (以下均指这种)，一般左逆满足 ①②③ 称为 正规化广义逆，一般右逆满足 ①②④ 称为 弱广义逆，只满足 ①② 的称为 自反广义逆

性质 广义逆唯一, $(\boldsymbol{A}^-)^-=\boldsymbol{A}$，$(\boldsymbol{A}^\dagger)^-=(\boldsymbol{A}^-)^\dagger\equiv\boldsymbol{A}^{-\dagger}$，对于宽矩阵 $\boldsymbol{A}^-=\boldsymbol{L}$，对于高矩阵, 列矢量 $\boldsymbol{A}^-=\boldsymbol{R}$

对 $\varphi\neq0$ 有 $(\varphi\boldsymbol{A})^-=\frac{1}{\varphi}\boldsymbol{A}^-$，$\boldsymbol{O}^-_{m\times n}=\boldsymbol{O}_{n\times m}$，$[\mathrm{diag}(a_{1\sim n})]^-=\mathrm{diag}(a^-_{1\sim n})$，若 $a_i=0$ 则 $a_i^-\equiv0$

若 $\boldsymbol{A}_i$ 相互正交 ($\boldsymbol{A}_i^\dagger\boldsymbol{A}_j=\boldsymbol{O}$ $(i\neq j)$) 则 $(\sum\boldsymbol{A}_i)^-=\sum\boldsymbol{A}_i^-$

$\mathrm{rank}\,\boldsymbol{A}=\mathrm{rank}\,\boldsymbol{A}^-=\mathrm{rank}(\boldsymbol{AA}^-)=\mathrm{rank}(\boldsymbol{A}^-\boldsymbol{A})=\mathrm{rank}(\boldsymbol{AA}^-\boldsymbol{A})=\mathrm{rank}(\boldsymbol{A}^-\boldsymbol{AA}^-)$

推论 $\boldsymbol{A}^{-\dagger}\boldsymbol{A}^\dagger\boldsymbol{A}=\boldsymbol{AA}^\dagger\boldsymbol{A}^{-\dagger}=\boldsymbol{A}$，$\boldsymbol{A}^\dagger\boldsymbol{A}^{-\dagger}\boldsymbol{A}^-=\boldsymbol{A}^-\boldsymbol{A}^{-\dagger}\boldsymbol{A}^\dagger=\boldsymbol{A}^-$，$\boldsymbol{A}^-\boldsymbol{AA}^\dagger=\boldsymbol{A}^\dagger\boldsymbol{AA}^-=\boldsymbol{A}^\dagger$，

$\boldsymbol{AA}^-\boldsymbol{A}^{-\dagger}=\boldsymbol{A}^{-\dagger}\boldsymbol{A}^-\boldsymbol{A}=\boldsymbol{A}^{-\dagger}$ (注意 $\boldsymbol{AA}^-$, $\boldsymbol{A}^-\boldsymbol{A}$, $\boldsymbol{A}^\dagger\boldsymbol{A}^{-\dagger}$, $\boldsymbol{A}^{-\dagger}\boldsymbol{A}^\dagger$ 都不等于 $\boldsymbol{I}$)

例 对称阵 $(\boldsymbol{A}^2)^-=(\boldsymbol{A}^-)^2$，厄米幂等矩阵 $\boldsymbol{A}^-=\boldsymbol{A}$

计算广义逆方法 ① 解矩阵方程 $\boldsymbol{AA}^\dagger\boldsymbol{X}^\dagger=\boldsymbol{A}$ 和 $\boldsymbol{A}^\dagger\boldsymbol{AY}=\boldsymbol{A}^\dagger$，然后 $\boldsymbol{A}^-=\boldsymbol{XAY}$〈数值〉

性质 若 $\boldsymbol{A}=\boldsymbol{BC}$，$\boldsymbol{B}$ 列满秩, $\boldsymbol{C}$ 行满秩, 则 $\boldsymbol{A}^-=\boldsymbol{C}^-\boldsymbol{B}^-=\boldsymbol{C}^\dagger(\boldsymbol{CC}^\dagger)^{-1}(\boldsymbol{B}^\dagger\boldsymbol{B})^{-1}\boldsymbol{B}^\dagger$

→ ② 满秩分解法, 化行简约阶梯型, 主元列组成 $\boldsymbol{B}$，阶梯型的非零行组成 $\boldsymbol{C}$

③ 若 $\boldsymbol{A}=\boldsymbol{U}\begin{bmatrix}\boldsymbol{\Sigma}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{bmatrix}\boldsymbol{V}^\dagger$，则 MP 广义逆 $\boldsymbol{A}^-=\boldsymbol{V}\begin{bmatrix}\boldsymbol{\Sigma}^{-1}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{bmatrix}\boldsymbol{U}^\dagger$

工程矩阵

「拉格朗日插值多项式〈数值〉设 $p(t)=x_0+x_1t+\cdots+x_{n-1}t^{n-1}$，测得 $t_{1\sim n},v_{1\sim n}$，要满足 $p(t_i)=v_i$，此即 $\boldsymbol{Vx}=\boldsymbol{v}$」

Vandermonde

范德蒙阵 行或列为等比数列的方阵 $\boldsymbol{V}_n$

「数学归纳法」$|\boldsymbol{V}|=\prod_{i<j}^{1\sim n}(t_i-t_j)$ → 所有 t 互异时非奇异

$$\begin{bmatrix}1&t_1&\dots&t_1^{n-1}\\1&t_2&\dots&t_2^{n-1}\\\vdots&\vdots&\ddots&\vdots\\1&t_n&\dots&t_n^{n-1}\end{bmatrix}$$

D Fourier T

记 $w=\mathrm{e}^{-\mathrm{i}2\pi/N}$ 为复单位根 $w^N=1$，则离散时间信号 $[x_{0\sim N-1}]^\mathsf{T}$ 的 离散傅氏变换 定义为 $y_k=\sum_{n=0}^{N-1}x_nw^{nk}$，$k=1\sim N-1$，矩阵形式 $\boldsymbol{y}=\boldsymbol{Fx}$，傅里叶阵 $\boldsymbol{F}_n=\big[w^{(i-1)(j-1)}\big]=\begin{bmatrix}1&1&\dots&1\\1&w&\dots&w^{N-1}\\\vdots&\vdots&\ddots&\vdots\\1&w^{N-1}&\dots&w^{(N-1)^2}\end{bmatrix}$

例 $\boldsymbol{F}_4=\begin{bmatrix}1&1&1&1\\1&\mathrm{i}&-1&-\mathrm{i}\\1&-1&1&-1\\1&-\mathrm{i}&-1&\mathrm{i}\end{bmatrix}$

性质 范德蒙, 对称, 幺正 → 其列称为 傅氏基

$\boldsymbol{F}^2=\begin{bmatrix}\boldsymbol{e}_1&\boldsymbol{e}_n&\boldsymbol{e}_{n-1}&\dots&\boldsymbol{e}_2\end{bmatrix}$ 是置换阵, $\boldsymbol{F}^4=\boldsymbol{I}$

离散傅氏逆变换 $\boldsymbol{F}^{-1}=\frac{1}{N}\boldsymbol{F}^\dagger=\frac{1}{N}\boldsymbol{F}^*$

FFT

快速傅氏变换 原来要做 n^2 次乘法, 现只需 $\frac{1}{2}n\,\mathrm{lb}\,n$

$\boldsymbol{F}_8=\begin{bmatrix}\boldsymbol{I}&\boldsymbol{D}\\\boldsymbol{I}&-\boldsymbol{D}\end{bmatrix}\begin{bmatrix}\boldsymbol{F}_4&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{F}_4\end{bmatrix}\boldsymbol{K}_{24}$，$\boldsymbol{D}=\mathrm{diag}(1,w,w^2,w^3)$，$\boldsymbol{K}_{24}=\begin{bmatrix}1&&&&&&&\\&&1&&&&&\\&&&&1&&&\\&&&&&&1&\\&1&&&&&&\\&&&1&&&&\\&&&&&1&&\\&&&&&&&1\end{bmatrix}$

Hadamard　　wavelets

哈达玛阵 $\boldsymbol{H}$ 只由 ±1 构成, $\frac{1}{\sqrt{n}}\boldsymbol{H}$ 为正交阵 → 其列称为 小波基 (Haar 1910)

$\boldsymbol{H}_2=\begin{bmatrix}1&1\\1&-1\end{bmatrix}$，「归纳法」$\boldsymbol{H}_{2^{k+1}}=\boldsymbol{H}_2\otimes\boldsymbol{H}_{2^k}$

(5 维不存在, 目前尚不知一般存在性)

性质 第一行第一列全 1，其它行列正负各半, 对称, 零迹

随机矩阵

summing vector

所有元素为 1 的矢量称为 求和矢量 $\mathbf{1}_n$ **例** 与自己内积 $\mathbf{1}_n^\mathsf{T}\mathbf{1}_n=n$，外积为 全 1 矩阵 $\mathbf{1}_m\mathbf{1}_n^\mathsf{T}=\mathbf{1}_{m\times n}$

例 $\mathbf{1}_{m\times n}\mathbf{1}_{n\times s}=n\mathbf{1}_{m\times s}$，$\mathbf{1}_{m\times n}\mathbf{1}_n=n\mathbf{1}_m$，$\mathbf{1}_m^\mathsf{T}\mathbf{1}_{m\times n}=m\mathbf{1}_n^\mathsf{T}$

对于方阵有 $\mathbf{1}_{n\times n}^2=n\mathbf{1}_{n\times n}$，故 $\bar{\mathbf{1}}_{n\times n}\equiv\frac{1}{n}\mathbf{1}_{n\times n}$ 为幂等阵, 本征值为 1 个 n 和 $(n-1)$ 个 0

centering matrix

中心化矩阵 $\boldsymbol{C}_n \equiv \boldsymbol{I}_n - \bar{\boldsymbol{1}}_{n\times n}$ **性质** 对称幂等阵, $\boldsymbol{C}_n\boldsymbol{1}_n=\boldsymbol{0}$, $\boldsymbol{C}_n\boldsymbol{1}_{n\times n}=\boldsymbol{1}_{n\times n}\boldsymbol{C}_n=\boldsymbol{0}$

设有数据矢量 $\boldsymbol{x}_{n\times 1}$, 则均值可表示为 $\overline{x}=\frac{1}{n}\boldsymbol{x}^\top\boldsymbol{1}_n=\frac{1}{n}\boldsymbol{1}_n^\top\boldsymbol{x}$, 而 $\boldsymbol{C}$ 的作用是减去均值 $\boldsymbol{C}\boldsymbol{x}=\boldsymbol{x}-\overline{x}\boldsymbol{1}_n$

求其内积, 发现中心化阵的二次型可表示协方差 $(\boldsymbol{C}\boldsymbol{x})^\top\boldsymbol{C}\boldsymbol{x}=\boldsymbol{x}^\top\boldsymbol{C}\boldsymbol{x}=\sum(x_i-\overline{x})^2$

设 x_i 为二阶矩有限的实或复随机变量, 期望 $\mu_i=\mathrm{E}(x_i)$ 则 $\boldsymbol{x}=[x_i]^\top$ 称为 **随机矢量** ,

$\boldsymbol{\mu}$ 称为 **均值矢量** , **自相关矩阵** $\boldsymbol{R}\equiv\mathrm{E}\left(\boldsymbol{x}\boldsymbol{x}^\dagger\right)$, 即 $r_{ij}=\mathrm{E}(x_ix_j^*)$ **性质** 厄米

自协 / 方差矩阵 $\mathrm{D}(\boldsymbol{x})=\mathrm{Cov}(\boldsymbol{x},\boldsymbol{x})\equiv\mathrm{E}\left((\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^\dagger\right)$, 即 $c_{ij}=\mathrm{E}\left((x_i-\mu_i)(x_j-\mu_j)^*\right)=\mathrm{E}\left(x_ix_j^*\right)-\mu_i\mu_j^*$

性质 厄米, 半正定, $\boldsymbol{C}_{\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}}=\boldsymbol{C}_{\boldsymbol{A}\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{C}_{\boldsymbol{x}}\boldsymbol{A}^\dagger$

互相关矩阵 $\boldsymbol{R}_{\boldsymbol{x},\boldsymbol{y}}\equiv\mathrm{E}\left(\boldsymbol{x}\boldsymbol{y}^\dagger\right)$ **互协方差矩阵** $\boldsymbol{C}_{\boldsymbol{x},\boldsymbol{y}}\equiv\mathrm{E}\left((\boldsymbol{x}-\boldsymbol{\mu}_x)(\boldsymbol{y}-\boldsymbol{\mu}_y)^\dagger\right)$ (可能非方阵, 即使是也不厄米)

性质 $\boldsymbol{C}_{\boldsymbol{x},\boldsymbol{y}}=\boldsymbol{C}_{\boldsymbol{y},\boldsymbol{x}}^\dagger$, 双线性, $\boldsymbol{C}_{\boldsymbol{A}\boldsymbol{x},\boldsymbol{B}\boldsymbol{y}}=\boldsymbol{A}\boldsymbol{C}_{\boldsymbol{x},\boldsymbol{y}}\boldsymbol{B}^\dagger$ **定理** $\boldsymbol{C}_{\boldsymbol{x},\boldsymbol{y}}=\boldsymbol{R}_{\boldsymbol{x},\boldsymbol{y}}-\boldsymbol{\mu}_x\boldsymbol{\mu}_y^\dagger$

correlation matrix

相关系数 $\rho_{12}\equiv\dfrac{c_{12}}{\sigma_1\sigma_2}$ **相关矩阵** $\left[c_{ij}/\sqrt{c_{ii}c_{jj}}\right]$ **性质** 半正定, 主对角元为 1 , 所有元绝对值 $\leqslant 1$

positive / probability vector / Markov / stochastic matrix

正矩阵 所有阵元为正实数, 记作 $\boldsymbol{A}>0$ **概率矢量** $\forall x_i\geqslant 0$, $\sum_i x_i=1$ **马尔可夫 / 随机矩阵** 非负, 方阵, 每列和为 1 $\to$ 1 是本征值 「$|\boldsymbol{A}-\boldsymbol{I}|\boldsymbol{1}=\boldsymbol{0}$ 故奇异」, 其它 $|\lambda_i|\leqslant 1$, 若每行和也为 1 称为 **双随机阵**

定理 $\boldsymbol{A}$ 是马尔可夫阵 $\Leftrightarrow$ $\boldsymbol{A}\boldsymbol{x}$ 是概率矢量 **例** 应用: 人口迁徙 $\boldsymbol{x}_{k+1}=\boldsymbol{A}\boldsymbol{x}_k$

性质 若 $\boldsymbol{A},\boldsymbol{B}$ 为马尔可夫阵, 则 $\boldsymbol{A}\boldsymbol{B}$ 也是, $p\boldsymbol{A}+(1-p)\boldsymbol{B}$ $(0\leqslant p\leqslant 1)$ 也是

矩阵微分

$\boldsymbol{x},\boldsymbol{X}$ 分别表示自变量为矢量/矩阵 (默认为列矢量) , $f,\boldsymbol{f},\boldsymbol{F}$ 分别表示函数值为数/矢量/矩阵

矩阵微分 $\mathrm{d}\boldsymbol{A}=\left[\mathrm{d}a_{ij}\right]$ **性质** 齐次叠加性, 转置, 迹可和微分换序, 积, 直积, 元素积的莱布尼茨法则同理

例 常数阵 $\mathrm{d}\boldsymbol{C}=\boldsymbol{O}$, $\mathrm{d}(\boldsymbol{C}\boldsymbol{X})=\boldsymbol{C}\,\mathrm{d}\boldsymbol{X}$

公式 $\mathrm{d}\ln\boldsymbol{X}=\boldsymbol{X}^{-1}\mathrm{d}\boldsymbol{X}$, $\mathrm{d}|\boldsymbol{X}|=|\boldsymbol{X}|\,\mathrm{tr}(\boldsymbol{X}^{-1}\mathrm{d}\boldsymbol{X})$, $\mathrm{d}(\boldsymbol{X}^{-1})=-\boldsymbol{X}^{-1}(\mathrm{d}\boldsymbol{X})\boldsymbol{X}^{-1}$

若 $\boldsymbol{X}$ 可逆, 则 $\boldsymbol{X}^{-1}$ 无穷次可微, $\mathrm{d}^k(\boldsymbol{X}^{-1})=(-1)^kk!(\boldsymbol{X}^{-1}\mathrm{d}\boldsymbol{X})^k\boldsymbol{X}^{-1}$, $k\in\mathbb{N}_+$

行矢量偏导 / 协变导数 $\partial_{\boldsymbol{x}^\top}=\left[\partial_{x_1}\ \cdots\ \partial_{x_n}\right]$ **列矢量偏导 / 逆变导数** $\partial_{\boldsymbol{x}}=\left[\partial_{x_1}\ \cdots\ \partial_{x_n}\right]^\top$

「 设有 n 个函数 $y_i=y_i(x_{1\sim n})$, $\mathrm{d}y_i(\boldsymbol{x})=\left[\partial_{x_1}y_i\ \cdots\ \partial_{x_n}y_i\right]\begin{bmatrix}\mathrm{d}x_1\\ \mathrm{d}x_n\end{bmatrix}=\partial_{\boldsymbol{x}^\top}y_i\,\mathrm{d}\boldsymbol{x}$ $\to$ $\mathrm{d}\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}$ 」

Jacobian

雅可比阵 $\boldsymbol{A}\equiv\left[\dfrac{\partial y_i}{\partial x_j}\right]$ 在定义域的每一点非奇异, 其行列式称为 **雅可比行列式**

性质 $\boldsymbol{A}_{(z/y)}\boldsymbol{A}_{(y/x)}=\boldsymbol{A}_{(z/x)}$ **例** $\boldsymbol{X},\boldsymbol{Y}$ 为对称方阵, $\boldsymbol{Y}=\boldsymbol{S}\boldsymbol{X}\boldsymbol{S}^\top$, 则 $|\boldsymbol{A}_{(Y/X)}|=|\boldsymbol{S}|^{n+1}$

辨识式 $\mathrm{d}f(\boldsymbol{x})=\partial_{\boldsymbol{x}^\top}f\,\mathrm{d}\boldsymbol{x}=\mathrm{tr}(\boldsymbol{A}\,\mathrm{d}\boldsymbol{x})$ $\Leftrightarrow$ $\partial_{\boldsymbol{x}^\top}f=\boldsymbol{A}$ $\Leftrightarrow$ $\partial_{\boldsymbol{x}}f=\boldsymbol{A}^\top$

函数	微分	雅可比阵	二阶微分	海森阵
$f(x):\mathbb{R}\to\mathbb{R}$	$\mathrm{d}f=A\,\mathrm{d}x$	$A\in\mathbb{R}$	$\mathrm{d}^2f=b(\mathrm{d}x)^2$	$b\in\mathbb{R}$
$f(\boldsymbol{x}):\mathbb{R}^m\to\mathbb{R}$	$\mathrm{d}f=\boldsymbol{A}\,\mathrm{d}\boldsymbol{x}$	$\boldsymbol{A}\in\mathbb{R}^{1\times m}$	$\mathrm{d}^2f=\mathrm{d}\boldsymbol{x}^\top\boldsymbol{B}\,\mathrm{d}\boldsymbol{x}$	$\frac{1}{2}(\boldsymbol{B}+\boldsymbol{B}^\top)\in\mathbb{R}^{m\times m}$
$f(\boldsymbol{X}):\mathbb{R}^{m\times n}\to\mathbb{R}$	$\mathrm{d}f=\mathrm{tr}(\boldsymbol{A}\,\mathrm{d}\boldsymbol{X})$	$\boldsymbol{A}\in\mathbb{R}^{n\times m}$	(上行 $\boldsymbol{x}$ 换成 $\mathrm{vec}\,\boldsymbol{X}$)	同上 $\in\mathbb{R}^{mn\times mn}$
$\boldsymbol{f}(\boldsymbol{x}):\mathbb{R}^m\to\mathbb{R}^p$	$\mathrm{d}\boldsymbol{f}=\boldsymbol{A}\,\mathrm{d}\boldsymbol{x}$	$\boldsymbol{A}\in\mathbb{R}^{p\times m}$	$\mathrm{d}^2\boldsymbol{f}=(\boldsymbol{I}_m\otimes\mathrm{d}\boldsymbol{x})^\top\boldsymbol{B}\,\mathrm{d}\boldsymbol{x}$	$\frac{1}{2}(\boldsymbol{B}+\boldsymbol{B}')\in\mathbb{R}^{pm\times m}$
$\boldsymbol{f}(\boldsymbol{X}):\mathbb{R}^{m\times n}\to\mathbb{R}^p$	$\mathrm{d}\boldsymbol{f}=\boldsymbol{A}\,\mathrm{d}(\mathrm{vec}\,\boldsymbol{X})$	$\boldsymbol{A}\in\mathbb{R}^{p\times mn}$	(上行 $\boldsymbol{x}$ 换成 $\mathrm{vec}\,\boldsymbol{X}$)	同上 $\in\mathbb{R}^{pmn\times mn}$
$\boldsymbol{F}(\boldsymbol{x}):\mathbb{R}^m\to\mathbb{R}^{p\times q}$	$\mathrm{d}(\mathrm{vec}\,\boldsymbol{F})=\boldsymbol{A}\,\mathrm{d}\boldsymbol{x}$	$\boldsymbol{A}\in\mathbb{R}^{pq\times m}$	(上上行 $\boldsymbol{f}$ 换成 $\mathrm{vec}\,\boldsymbol{F}$)	同上 $\in\mathbb{R}^{pmq\times m}$
$\boldsymbol{F}(\boldsymbol{X}):\mathbb{R}^{m\times n}\to\mathbb{R}^{p\times q}$	$\mathrm{d}\boldsymbol{F}=\boldsymbol{A}(\mathrm{d}\boldsymbol{X})\boldsymbol{B}+\boldsymbol{C}(\mathrm{d}\boldsymbol{X}^\top)\boldsymbol{D}$	$(\boldsymbol{B}^\top\otimes\boldsymbol{A}+(\boldsymbol{D}^\top\otimes\boldsymbol{C})\boldsymbol{K}_{mn})\in\mathbb{R}^{pq\times mn}$	(上行 $\boldsymbol{x}$ 换成 $\mathrm{vec}\,\boldsymbol{X}$)	同上 $\in\mathbb{R}^{pmqn\times mn}$

其中 $\boldsymbol{B}'$ 仍为 $pmn\times mn$ 列分块阵, 对每个分块做转置, $\boldsymbol{F}$ 行的 $\boldsymbol{B}'$ 是 $pmqn\times mn$

例 二次型函数 $f(\boldsymbol{x})=\boldsymbol{x}^\top\boldsymbol{G}\boldsymbol{x}$, $\boldsymbol{G}$ 为常数方阵, 由 $\mathrm{d}f=\mathrm{d}(\mathrm{tr}(\boldsymbol{x}^\top\boldsymbol{G}\boldsymbol{x}))=\mathrm{tr}(\mathrm{d}\boldsymbol{x}^\top\boldsymbol{G}\boldsymbol{x}+\boldsymbol{x}^\top\boldsymbol{G}\,\mathrm{d}\boldsymbol{x})$ 前者转置不改变迹 $=\mathrm{tr}(\boldsymbol{x}^\top\boldsymbol{G}^\top\mathrm{d}\boldsymbol{x}+\boldsymbol{x}^\top\boldsymbol{G}\,\mathrm{d}\boldsymbol{x})$ $\to$ $\partial_{\boldsymbol{x}}f=(\boldsymbol{x}^\top(\boldsymbol{G}^\top+\boldsymbol{G}))^\top=(\boldsymbol{G}^\top+\boldsymbol{G})\boldsymbol{x}$

「标量函数总可以写成迹函数的形式, 因为 $f(\boldsymbol{X})=\mathrm{tr}(f(\boldsymbol{X}))$, 如果有多项相乘总可换到最右端 $\mathrm{tr}(\boldsymbol{A}(\mathrm{d}\boldsymbol{X})\boldsymbol{B})=\mathrm{tr}(\boldsymbol{B}\boldsymbol{A}\,\mathrm{d}\boldsymbol{X})$, 如果是 $\mathrm{d}\boldsymbol{X}^\top$ 总可以转置」

迹函数 **例** $\mathrm{d\,tr}(XAXB)=\mathrm{tr}(\mathrm{d}(XAXB))=\mathrm{tr}((\mathrm{d}X)AXB+XA(\mathrm{d}X)B)=\mathrm{tr}((AXB+BXA)\mathrm{d}X)$

f	$\mathrm{tr}(CX)$	$\mathrm{tr}(X^2)$	$\mathrm{tr}(AX^{-1}B)$	$\mathrm{tr}(XAX^\top B)$
$\mathrm{d}f$	$\mathrm{tr}(C\,\mathrm{d}X)$	$2\,\mathrm{tr}(X\,\mathrm{d}X)$	$-\mathrm{tr}(X^{-1}BAX^{-1}\mathrm{d}X)$	$\mathrm{tr}((AX^\top B+(BXA)^\top)\mathrm{d}X)$

行列式函数微分表 **例** $\mathrm{d}\ln|X|=|X|^{-1}\mathrm{d}|X|=|X|^{-1}|X|\,\mathrm{tr}(X^{-1}\mathrm{d}X)=\mathrm{tr}(X^{-1}\mathrm{d}X)$

f	$\lvert X^k\rvert$, $k\in\mathbb{Z}$	$\lvert AXB\rvert$	$\lvert Q\rvert$, $Q=X^\top AX$
$\mathrm{d}f$	$k\lvert X^k\rvert\,\mathrm{tr}(X^{-1}\mathrm{d}X)$	$\lvert AXB\rvert\,\mathrm{tr}(B(AXB)^{-1}BAA\,\mathrm{d}X)$	$\lvert Q\rvert\,\mathrm{tr}((Q^{-\top}X^\top A^\top+Q^{-1}X^\top A)\mathrm{d}X)$

矩阵函数 **例** $F=X^\top X$, $\mathrm{d}F=X^\top\mathrm{d}X+\mathrm{d}X^\top X$, 和辨识式对比知雅可比阵为 $I_n\otimes X^\top+(X^\top\otimes I_n)K_{mn}$

F	X^{-1}	X^k	$\ln(X)$	$\exp(X)$
$\mathrm{d}F$	$-X^{-1}(\mathrm{d}X)X^{-1}$	$\sum_{i=1}^k X^{i-1}(\mathrm{d}X)X^{k-i}$	$X^{-1}\mathrm{d}X$	$\sum_{k=0}^\infty \frac{1}{(k+1)!}\sum_{i=0}^k X^i(\mathrm{d}X)X^{k-i}$

凸优化

〈微积分〉「多元函数泰勒展开 $f(y)=f(x)+\sum_i(y_i-x_i)f'_{x_i}|_x+\frac{1}{2}\sum_{ij}(y_i-x_i)(y_j-x_j)f''_{x_ix_j}|_x+\dots$ 」

临界点 一阶偏导为零 $\partial_x f|_{x=c}=0$, 在 c 附近 $f(c+\Delta x)\approx f(c)+\frac{1}{2}(\Delta x)^\top H\Delta x$

$\rightarrow$ 存在极小/大值 $\Leftrightarrow$ 二阶导数阵 H 正/负定, 若 $x^\top Hx$ 在 $\mathcal{D}$ 所有点非负, 则称 f 是 $\mathcal{D}$ 上的 凸函数

Hessian
海森阵 $H\equiv[\partial_{x_i}\partial_{x_j}f]=\partial_x(\partial_{x^\top}f)\equiv\nabla_x^2 f$ **性质** 对称阵「偏导可换序」

(用二阶微分矩阵求海森较简单, 因 $\mathrm{d}x$ 不是矢量 x 的函数, 故 $\mathrm{d}^2x=0$, $\mathrm{d}^2f=(\mathrm{d}x)^\top(\partial_x\partial_{x^\top}\mathrm{d}f)\mathrm{d}x$)

记 $\mathrm{d}f=A\,\mathrm{d}x$ 中的 $A=(\mathrm{d}x)^\top B$, 则 $\mathrm{d}^2f=\mathrm{d}(A\,\mathrm{d}x)=(\mathrm{d}x)^\top B\,\mathrm{d}x$, 对比得 $H_xf=B$, 为确保实对称, 取 $H=\frac{1}{2}(B^\top+B)$ **例** $\mathrm{tr}(X^2)$ 的海森阵为 $2K_{nn}$

系数非负且和为 1 的线性组合称为 凸组合 , 从几何中任取一组元素任意凸组合仍在集合内称为 凸集

$\Leftrightarrow$ 连接集合任意两点的线段仍在集合中〈拓扑〉

convex
目标函数为凸函数, 且定义域为凸集的优化问题, 称为 凸优化

(若有不等式约束, 需为凸函数, 若有等式约束, 需为仿射函数)

定理 无约束凸函数 $f(x)$ 的任何极小点就是全局极小点, 若可微则求解 $\partial_x f=0$ 可得该极小点

平滑凸优化的一阶算法 $x_{k+1}=x_k+l_k\Delta x_k$, l_k 为步长, Δx_k 称为步行方向

descent method
梯度下降法 迭代每一步选负梯度方向爬 $\Delta x_k=-\nabla f(x_k)$