矩阵分析笔记

1 矩阵的基本概念

• Hermitian 矩阵：$\boldsymbol{A^H}=\boldsymbol{A}$。幂等矩阵：$\boldsymbol{A}^2=\boldsymbol{A}$。

• 内积：$\langle\boldsymbol{x},\boldsymbol{y}\rangle=\boldsymbol{x}^H\boldsymbol{y}$，$\langle\boldsymbol{x}(t),\boldsymbol{y}(t)\rangle=\int\boldsymbol{x}^H\boldsymbol{y}dt$，$\langle\boldsymbol{x}(\xi),\boldsymbol{y}(\xi)\rangle=E\{\boldsymbol{x}^H\boldsymbol{y}\}$

• L_p范数：$\|\boldsymbol{x}\|_p=(\sum|x_i|^p)^{1/p}$，距离$d(\boldsymbol{x},\boldsymbol{y})=\|\boldsymbol{x}-\boldsymbol{y}\|$。

• 范数性质（P27）

• 向量夹角：$\cos\theta=\frac{\langle x,y\rangle}{\|x\|\|y\|}$。常数向量、函数向量正交充要条件是内积为零，随机向量正价条件是外积为零。

• 均值向量$\boldsymbol{\mu_x}=E\{\boldsymbol{x}(\xi)\}$，自相关矩阵$\boldsymbol{R_x}=E\{\boldsymbol{x}(\xi)\boldsymbol{x}^H(\xi)\}$，自协方差矩阵$\boldsymbol{C_x}=E\{[\boldsymbol{x}(\xi)-\boldsymbol{\mu_x}][\boldsymbol{x}(\xi)-\boldsymbol{\mu_x}]^H\}=\boldsymbol{R_x}-\boldsymbol{\mu_x}\boldsymbol{\mu_x}^H$。互相关矩阵$\boldsymbol{R_{xy}}=E\{\boldsymbol{x}(\xi)\boldsymbol{y}^H(\xi)\}$，互协方差矩阵$\boldsymbol{C_{xy}}=E\{[\boldsymbol{x}(\xi)-\boldsymbol{\mu_x}][\boldsymbol{y}(\xi)-\boldsymbol{\mu_y}]^H\}=\boldsymbol{R_{xy}}-\boldsymbol{\mu_x}\boldsymbol{\mu_y}^H$。相关系数$\rho_{xy}=\frac{c_{xy}}{\sigma_x\sigma_y}$。统计不相关充要条件为互协方差矩阵为零，正交充要条件为互相关矩阵为零。

• $\boldsymbol{x}(\xi)\sim\mathcal{N}(\boldsymbol{\mu_x},C_{\boldsymbol{x}})$，若 x_i 独立同分布，$f(\boldsymbol{x})=\frac{1}{(2\pi)^{n/2}|C_{\boldsymbol{x}}|^{1/2}}exp(-\frac{1}{2}(\boldsymbol{x}(\xi)-\boldsymbol{\mu_x})^TC_x^{-1}(\boldsymbol{x}(\xi)-\boldsymbol{\mu_x}))$。

•欧氏距离：$D_E(\boldsymbol{x},\boldsymbol{y})=\|\boldsymbol{x}-\boldsymbol{y}\|_2$；马氏距离：$D_M(\boldsymbol{x},\boldsymbol{y})=\sqrt{(\boldsymbol{x}-\boldsymbol{y})^TC_x^{-1}(\boldsymbol{x}-\boldsymbol{y})}$。

• 矩阵的L_p范数：$\|\boldsymbol{A}\|_p=(\sum_{ij}|a_{ij}|^p)^{1/p}$。

• 矩阵范数性质（P33），内积与范数关系（P35）

• 矩阵的行列式：$det(\boldsymbol{A})=\sum_j a_{ij}(-1)^{i+j}det(\boldsymbol{A_{ij}})$。

• 行列式性质、特征值及性质、迹及性质、秩及性质、逆及性质（P47~P55）

• 矩阵和的求逆公式：$(\boldsymbol{A}+\boldsymbol{UBV})^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{U}(\boldsymbol{I}+\boldsymbol{BVA}^{-1}\boldsymbol{U})^{-1}\boldsymbol{BVA}^{-1}$。

• 分块矩阵的逆（P57），Hermitian 矩阵求逆引理（P58~59）

• 左逆矩阵：$\boldsymbol{LA}=\boldsymbol{I},m\geq n$；右逆矩阵：$\boldsymbol{AR}=\boldsymbol{I},m\leq n$。若$rank(\boldsymbol{A})=n$，左伪

逆矩阵$\boldsymbol{L}=(\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\boldsymbol{H}}$，若$rank(\boldsymbol{A})=m$，右伪逆矩阵$\boldsymbol{R}=\boldsymbol{A}^{\boldsymbol{H}}(\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}})^{-1}$，伪逆矩阵唯一确定。

• 广义逆矩阵 Moore-Penrose 条件：$\boldsymbol{AGA}=\boldsymbol{A},\boldsymbol{GAG}=\boldsymbol{G},\boldsymbol{AG}$和$\boldsymbol{GA}$为 Hermitian 矩阵。

• M-P 矩阵计算（P64~P67）

• Kronecker 积：$\boldsymbol{A}\otimes\boldsymbol{B}=\left[a_{ij}B\right],[\boldsymbol{A}\otimes\boldsymbol{B}]_{left}=\left[\boldsymbol{A}b_{ij}\right]$

• 直和：$\boldsymbol{A}\oplus\boldsymbol{B}=\begin{bmatrix}\boldsymbol{A} & \\ & \boldsymbol{B}\end{bmatrix}$

• Hadamard 积：$[\boldsymbol{A}*\boldsymbol{B}]_{ij}=a_{ij}b_{ij}$，若$\boldsymbol{A},\boldsymbol{B}$正定（半正定），则$\boldsymbol{A}*\boldsymbol{B}$正定（半正定）

• 直和与 Hadamard 积性质（P67~P70）

2 特殊矩阵

• 基本矩阵$\boldsymbol{E}_{mn}=e_me_n^T$。

• 基本矩阵性质（P101）

• Ⅰ型基本矩阵$\boldsymbol{E}_{(p,q)}$，Ⅱ型基本矩阵$\boldsymbol{E}_{\alpha(p)}$，Ⅲ型基本矩阵$\boldsymbol{E}_{(p)+\alpha(q)}$，对矩阵左乘表示行操作，右乘表示列操作。

• 置换矩阵：各行列有且仅有一个非零元素 1。移位矩阵$\boldsymbol{P}=(e_2,\cdots,e_n,e_1)^T$。互换矩阵$\boldsymbol{J}=(e_n,\cdots,e_1)^T$。选择矩阵$\boldsymbol{P}_{(M-1)\times M}=[I_{M-1},0]$。交换矩阵$\boldsymbol{K}_{mn}vec(\boldsymbol{A})=vec(\boldsymbol{A}^T)$。广义置换矩阵：各行列有且仅有一个非零元素。

• 酉矩阵：$\boldsymbol{U}\boldsymbol{U}^{\boldsymbol{H}}=\boldsymbol{U}^{\boldsymbol{H}}\boldsymbol{U}=\boldsymbol{I}$。仿酉矩阵：只满足$\boldsymbol{U}\boldsymbol{U}^{\boldsymbol{H}}=\boldsymbol{I}$或$\boldsymbol{U}^{\boldsymbol{H}}\boldsymbol{U}=\boldsymbol{I}$。正交矩阵与半正交矩阵定义与之类似。

• 酉矩阵性质（P111）

• 正规矩阵：$\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{H}}=\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{A}$。

•求和向量$\mathbf{1}_n=[1,\cdots,1]^T$，$J_{p\times q}=\mathbf{1}_p\mathbf{1}_q^T$，中心化矩阵$\boldsymbol{C}_{\boldsymbol{n}}=\boldsymbol{I}_{\boldsymbol{n}}-\frac{1}{n}\boldsymbol{J}_{\boldsymbol{n}}$，$\sum(x_i-\bar{x})^2=$

$\boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x}$。

• $\boldsymbol{B}\sim\boldsymbol{A}$：$\boldsymbol{B}=\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$，$S$非奇异。

• 相似矩阵性质（P117~P118）

• 相合矩阵：$\boldsymbol{B}=\boldsymbol{C}^{\boldsymbol{H}}\boldsymbol{A}\boldsymbol{C}$，$\boldsymbol{C}$非奇异。相合矩阵的二次型相同。

• 三角矩阵性质（P114）

• Vandermonde 矩阵$\boldsymbol{A}=\begin{bmatrix} a_1^0 & \cdots & a_N^0 \\ \vdots & \ddots & \vdots \\ a_1^{N-1} & \cdots & a_N^{N-1} \end{bmatrix}$，$det(\boldsymbol{A})=\prod_{i>j}(a_i-a_j)$

• Fourier 矩阵$\boldsymbol{F}_{ik}=\omega^{(i-1)(k-1)},\omega=e^{-j2\pi/N}$

• Fourier 矩阵性质（P124）

• Hadamard 矩阵：所有元素取±1，且$\boldsymbol{H_n}\boldsymbol{H_n^T}=\boldsymbol{H_n^T}\boldsymbol{H_n}=n\boldsymbol{I_n}$。

• Hadamard 矩阵性质（P129~P130）

• Toeplitz 矩阵$\boldsymbol{A}_{mn}=a_{m-n}$。对称 Toeplitz 矩阵：$a_{-i}=a_i$。Hermitian Toeplitz 矩阵：$a_{-i}=a_i^*$。

3 矩阵微分

• 根据函数$f(z)$在$z=0$点的幂级数构造矩阵函数$f(\boldsymbol{A})=\sum_k c_k\boldsymbol{A}^k$。

• 若$\boldsymbol{AB}=\boldsymbol{BA}$，则$e^{\boldsymbol{A}}e^{\boldsymbol{B}}=e^{\boldsymbol{B}}e^{\boldsymbol{A}}=e^{\boldsymbol{A}+\boldsymbol{B}}$。

• 若 $\boldsymbol{AB}=\boldsymbol{BA}$ ，则 $cos(\boldsymbol{A}+\boldsymbol{B})=cos\boldsymbol{A}cos\boldsymbol{B}-sin\boldsymbol{A}sin\boldsymbol{B},sin(\boldsymbol{A}+\boldsymbol{B})=sin\boldsymbol{A}cos\boldsymbol{B}+cos\boldsymbol{A}sin\boldsymbol{B}$。

• 设$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$，则$f(\boldsymbol{A})=\boldsymbol{P}\begin{bmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{bmatrix}\boldsymbol{P}^{-1}$。

• 偏导算子$D_{\boldsymbol{x}}=\frac{\partial}{\partial\boldsymbol{x}^T},D_{vec(\boldsymbol{X})}=\frac{\partial}{\partial vec^T(\boldsymbol{X})},D_{\boldsymbol{X}}=\frac{\partial}{\partial\boldsymbol{X}^T}$，梯度算子$\nabla_{\boldsymbol{x}}=\frac{\partial}{\partial\boldsymbol{x}},\nabla_{vec(\boldsymbol{X})}=\frac{\partial}{\partial vec(\boldsymbol{X})}$，$rvec(D_{\boldsymbol{X}})=D_{vec(\boldsymbol{X})},\nabla_{\boldsymbol{x}}=D_{\boldsymbol{x}}^T$。

• $[D_{\boldsymbol{x}}\boldsymbol{f}(\boldsymbol{x})]_{ij}=\frac{\partial f_i(\boldsymbol{x})}{\partial x_j},D_{\boldsymbol{X}}\boldsymbol{F}(\boldsymbol{X})=\frac{\partial vec(\boldsymbol{F}(\boldsymbol{X}))}{\partial vec^T(\boldsymbol{X})}$

• $df(\boldsymbol{x}) = tr(\boldsymbol{A}d\boldsymbol{x}), \boldsymbol{A} = D_{\boldsymbol{x}}f(\boldsymbol{x}); df(\boldsymbol{X}) = tr(\boldsymbol{A}d\boldsymbol{X}), \boldsymbol{A} = D_{\boldsymbol{X}}f(\boldsymbol{X})$

• 矩阵微分的性质和常用公式（P152~P153,P158,P160）

• $d(vec\boldsymbol{F}(\boldsymbol{X})) = \boldsymbol{A}d(vec\boldsymbol{X}) + \boldsymbol{B}d(vec\boldsymbol{X}^T) = (\boldsymbol{A} + \boldsymbol{B}\boldsymbol{K}_{\boldsymbol{mn}})d(vec\boldsymbol{X}), D_{\boldsymbol{X}}\boldsymbol{F}(\boldsymbol{X}) = \boldsymbol{A} + \boldsymbol{B}\boldsymbol{K}_{\boldsymbol{mn}}$； $d\boldsymbol{F}(\boldsymbol{X}) = \boldsymbol{A}(d\boldsymbol{X})\boldsymbol{B} + \boldsymbol{C}(d\boldsymbol{X}^T)\boldsymbol{D}, D_{\boldsymbol{X}}\boldsymbol{F}(\boldsymbol{X}) = (\boldsymbol{B}^{\boldsymbol{T}} \otimes \boldsymbol{A}) + (\boldsymbol{D}^{\boldsymbol{T}} \otimes \boldsymbol{C})\boldsymbol{K}_{\boldsymbol{mn}}$。

• $\boldsymbol{H}[f(\boldsymbol{x})] = \nabla_{\boldsymbol{x}}^2 f(\boldsymbol{x}) = \nabla_{\boldsymbol{x}}(D_{\boldsymbol{x}}f(\boldsymbol{x})), \boldsymbol{H}^T[f(\boldsymbol{x})] = \boldsymbol{H}[f(\boldsymbol{x})]$

• $d^2f(\boldsymbol{x}) = (d\boldsymbol{x})^T\boldsymbol{B}d\boldsymbol{x}, \boldsymbol{H}[f(\boldsymbol{x})] = \frac{1}{2}(\boldsymbol{B}^T + \boldsymbol{B}); d^2f(\boldsymbol{X}) = (d(vec(\boldsymbol{X})))^T\boldsymbol{B}d(vec\boldsymbol{X}), \boldsymbol{H}[f(\boldsymbol{x})] = \frac{1}{2}(\boldsymbol{B}^T + \boldsymbol{B})$

• $\frac{\partial}{\partial z} = \frac{1}{2}(\frac{\partial}{\partial x} - j\frac{\partial}{\partial y}), \frac{\partial}{\partial z^*} = \frac{1}{2}(\frac{\partial}{\partial x} + j\frac{\partial}{\partial y})$

• 复变函数偏导的常用公式、法则（P173~P174）

• 协梯度算子$D_{\boldsymbol{z}} = \frac{\partial}{\partial \boldsymbol{z}^T}$，共轭协梯度算子$D_{\boldsymbol{z}^*} = \frac{\partial}{\partial \boldsymbol{z}^H}$，梯度算子$\nabla_{\boldsymbol{z}} = \frac{\partial}{\partial \boldsymbol{z}}$，共轭梯度算子$\nabla_{\boldsymbol{z}^*} = \frac{\partial}{\partial \boldsymbol{z}^*}$

• 复偏导算子重要结果（P177）

• 梯度矩阵运算法则（P179）

• $df(\boldsymbol{Z}, \boldsymbol{Z}^*) = tr(\boldsymbol{A}d\boldsymbol{Z} + \boldsymbol{B}d\boldsymbol{Z}^*), \nabla_{\boldsymbol{Z}}f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{A}^T, \nabla_{\boldsymbol{Z}^*}f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \boldsymbol{B}^T$

• 几种迹函数、行列式函数的复矩阵微分（P183~P184）

4 无约束优化和凸优化基础

• $f(\boldsymbol{x})$局部极小点：$\nabla_{\boldsymbol{x}}f(x) = 0, \nabla_{\boldsymbol{x}}^2 f(\boldsymbol{x}) \succcurlyeq 0$。$f(\boldsymbol{z})$局部极小点：$\nabla_{\boldsymbol{z}^*}f(\boldsymbol{z}, \boldsymbol{z}^*) = 0, \boldsymbol{H}f(\boldsymbol{z}, \boldsymbol{z}^*) \succcurlyeq 0$。凸函数的局部极小点即为全局极小点。

• 凸函数、强凸函数、拟凸函数定义（P213~P214）

• 一般优化问题：$\min_{\boldsymbol{x} \in \mathcal{D}} f_0(\boldsymbol{x}) \ s.t. f_i(\boldsymbol{x}) \leq b_i, i = 1, \cdots, m$。如果目标函数和约束函数都是凸函数，则为凸优化问题。

• 凸集定义与运算性质（P211~P212）

• 函数$f(\boldsymbol{x})$为凸函数的充要条件是$g(t)=f(\boldsymbol{x}+t\boldsymbol{v})$对任意$\boldsymbol{x},\boldsymbol{v}$都是凸函数。

• 函数$f(\boldsymbol{x})$为凸函数的充要条件是$f(\boldsymbol{y})\geq f(\boldsymbol{x})+\nabla f^T(\boldsymbol{x})(\boldsymbol{y}-\boldsymbol{x})$。

• 设凸优化问题最优解p^*，记$L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{v})=f_0(\boldsymbol{x})+\sum\lambda_i f_i(\boldsymbol{x})+\sum v_i h_i(\boldsymbol{x})$，拉格朗日对偶函数$g(\boldsymbol{\lambda},\boldsymbol{v})=\inf_{\boldsymbol{x}} L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{v})$是凹函数。对偶问题：$\max_{\boldsymbol{\lambda}\succcurlyeq 0,\boldsymbol{v}} J_D(\boldsymbol{\lambda},\boldsymbol{v})$，其最优解$d^*\leq p^*$。

• 拉格朗日对偶无约束优化问题 KKT 条件：$\begin{cases} f_i(\boldsymbol{x})\leq 0 \\ h_i(\boldsymbol{x})=0 \\ \boldsymbol{\lambda}\succcurlyeq 0 \\ \lambda_i f_i(\boldsymbol{x})=0 \\ \nabla L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{v})=0 \end{cases}$。

• 凸优化与拉格朗日对偶无约束优化强对偶。非凸优化，解在可行域内域则强对偶（Slater 定理），否则弱对偶。

5 奇异值分解

• 条件数（P287）

• 对矩阵$\boldsymbol{A}$，存在酉矩阵$\boldsymbol{U},\boldsymbol{V}$，使得$\boldsymbol{A}=\boldsymbol{U}\sum\boldsymbol{V}^{\boldsymbol{H}},\sum=\begin{bmatrix}\sum_{\boldsymbol{1}} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}$，其中$\sum_{\boldsymbol{1}}=diag(\sigma_1,\cdots,\sigma_r),\sigma_1\geq\cdots\geq\sigma_r>0,r=rank(\boldsymbol{A})$。$\sigma_1,\cdots,\sigma_r$以及$\sigma_{r+1}=\cdots=\sigma_n=0$称为$\boldsymbol{A}$的奇异值。

• 奇异值的分解形式（P289~P290）

• 对优化问题$\min_{rank(\boldsymbol{B})=k<rank(\boldsymbol{A})}\|\boldsymbol{A}-\boldsymbol{B}\|_F^2$，解为$\boldsymbol{A}_{\boldsymbol{k}}=\sum_{i=1}^{k}\sigma_i\boldsymbol{u}_i\boldsymbol{v}_i^H$，且$\|\boldsymbol{A}-\boldsymbol{A}_{\boldsymbol{k}}\|_F^2=\sum_{i=k+1}^{r}\sigma_i^2$。

• 奇异值分解的唯一性（P291~P292）

• 奇异值的性质（P292~P295）

• 奇异值分解的应用（P301~P304）

• 乘积奇异值分解（P298~P299）

• 广义奇异值分解（P304~P306）

6 最小二乘

• 最小二乘问题：$\min_{\boldsymbol{x}} \|\boldsymbol{Ax}-\boldsymbol{b}\|_p$。$p=1$ 时称为最小绝对偏差问题，$p=2$ 称为线性最小二乘问题，$p=\infty$称为最小最大问题。

• Guass-Markov 定理：对于$\boldsymbol{Ax}=\boldsymbol{b}+\boldsymbol{e}$，当且仅当$\boldsymbol{A}$列满秩时，存在最优无偏解，它由最小二乘解$\hat{\boldsymbol{x}}_{\boldsymbol{SL}}=(\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{b}$给出，具有最小方差。

• Tikhonov 正则化：对秩亏损情况，$\hat{\boldsymbol{x}}_{\boldsymbol{Tik}}=(\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{A}+\boldsymbol{\lambda I})^{-1}\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{b}$。

• 普通最小二乘（b 有误差），解：$\hat{\boldsymbol{x}}_{\boldsymbol{SL}}=(\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{b}$

• 数据最小二乘（A 有误差），解：$\hat{\boldsymbol{x}}_{\boldsymbol{DLS}}=\min_{\boldsymbol{x}}\frac{(\boldsymbol{Ax}-\boldsymbol{b})^{\boldsymbol{H}}(\boldsymbol{Ax}-\boldsymbol{b})}{\boldsymbol{x}^{\boldsymbol{H}}\boldsymbol{x}}$。

• 总体最小二乘解：$\min_{\boldsymbol{x}}\frac{(\boldsymbol{Ax}-\boldsymbol{b})^{\boldsymbol{H}}(\boldsymbol{Ax}-\boldsymbol{b})}{1+\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{x}}$。（P337~P348）

7 特征分析

• 代数多重度、几何多重度（P402）

• 特征值和特征向量性质（P408~P410）

• 奇异值分解和特征值分解的区别（P403）

• 若实矩阵$\boldsymbol{A}$具有 n 个线性无关的特征向量（或 n 个不同的特征值），则其可对角化。

•Cayley-Hamilton 定理：特征多项式$p(x)=det(\boldsymbol{A}-x\boldsymbol{I})$是使方阵$\boldsymbol{A}$零化的多项式。

• 逆矩阵的计算（P417）

• 矩阵幂的计算（P419）

•迷向圆变换$\boldsymbol{y}=\sum_{\boldsymbol{x}}^{-1/2}\boldsymbol{U}_x^H(\boldsymbol{x}-\boldsymbol{m}_{\boldsymbol{x}})$，其中$\boldsymbol{x}$为随机向量，协方差矩阵$\boldsymbol{C}_{\boldsymbol{x}}=\boldsymbol{U}_{\boldsymbol{x}}\sum_{\boldsymbol{x}}\boldsymbol{U}_x^H$。$\boldsymbol{y}$的所有分量都是零均值、具有单位方差的统计不相关随机变量。

• 离散 Karhunen-Loeve 变换（P428~P430）

• 满足$\boldsymbol{Au}=\lambda\boldsymbol{Bu}$的标量$\lambda$和非零向量$\boldsymbol{u}$称为矩阵束$(\boldsymbol{A},\boldsymbol{B})$的广义特征值和广义特

征向量，$\lambda(\boldsymbol{A},\boldsymbol{B})=\{z:det(\boldsymbol{A}-z\boldsymbol{B})=0\}$。

• 广义特征值性质（P434）

• $(\boldsymbol{XAY},\boldsymbol{XBY})$是$(\boldsymbol{A},\boldsymbol{B})$的等价矩阵束，$\boldsymbol{X},\boldsymbol{Y}$非奇异。

• 广义特征值分解的总体最小二乘方法（P437）

• Hermitian 矩阵$\boldsymbol{A}$Rayleigh 商：$R(\boldsymbol{x})=\frac{\boldsymbol{x}^{\boldsymbol{H}}\boldsymbol{Ax}}{\boldsymbol{x}^{\boldsymbol{H}}\boldsymbol{x}}$，$\boldsymbol{x}$可限定于单位球使得$R(\boldsymbol{x})$取最值。

• Rayleigh 商的性质（P443）

• 广义 Rayleigh 商（P447~P448）

• 复矩阵$\boldsymbol{A}$酉相似于对角矩阵的充要条件是$\boldsymbol{A}$为正规矩阵

• 复矩阵$\boldsymbol{A}$相似于对角矩阵的充要条件是$\boldsymbol{A}$有 n 个线性无关的特征向量

• 存在可逆矩阵$\boldsymbol{P}$，使任意矩阵$\boldsymbol{A}$相似于块对角矩阵，$J_i(\lambda_i)=\begin{pmatrix}\lambda_i & 1 & \\ & \ddots & 1 \\ & & \lambda_i\end{pmatrix}$

8 子空间

• 线性空间要求加法满足结合律、交换律、零元素、存在负元素，乘法满足数因子分配律、分配律、结合律、单位元。

•线性空间$\mathcal{V}$中线性无关向量组所含向量的最大个数称为$\mathcal{V}$的维数，记作$dim\,\mathcal{V}=n$。

•若向量集$S=\{\boldsymbol{u_1},\cdots,\boldsymbol{u_m}\}$是向量空间的子集，则$\mathcal{W}=Span\{\boldsymbol{u_1},\cdots,\boldsymbol{u_m}\}=\{\boldsymbol{u}:\boldsymbol{u}=a_1\boldsymbol{u_1}+\cdots+a_m\boldsymbol{u_m}\}$称为由$\boldsymbol{u_1},\cdots,\boldsymbol{u_m}$张成的子空间。$\boldsymbol{u_1},\cdots,\boldsymbol{u_m}$称为$\mathcal{W}$的生成元，$S$称为$\mathcal{W}$的张成集。

• 张成集定理：（1）若S中某向量是其他向量的线性组合，删除后仍张成$\mathcal{W}$；（2）若$\mathcal{W}$非平凡，则在S中存在某线性无关向量的子集合，它张成$\mathcal{W}$。

• $\mathcal{W}$的子集合$\{\boldsymbol{u_1},\cdots,\boldsymbol{u_p}\}$称为$\mathcal{W}$的一组基，如果$\mathcal{W}=Span\{\boldsymbol{u_1},\cdots,\boldsymbol{u_p}\}$，$\boldsymbol{u_1},\cdots,\boldsymbol{u_p}$线性无关。

• 基表示定理：令$S=\{\boldsymbol{u_1},\cdots,\boldsymbol{u_m}\}$是$\mathcal{W}$的一组基，则$\mathcal{W}$中任意向量都可以表示为$\boldsymbol{u_1},\cdots,\boldsymbol{u_m}$的线性组合。

• 若U和V是$\mathcal{W}$的两组不同的基，若基变换$V=UC$，称C为U到V的过渡矩阵。

• 设$\boldsymbol{x}\in\mathcal{W}$在$U$和$V$下坐标为$\boldsymbol{\xi},\boldsymbol{\eta}$，那么$\boldsymbol{\xi}=C^{-1}\boldsymbol{\eta}$。

• 子空间$S=S_1\cap\cdots\cap S_n$表示$S_1,\cdots,S_n$共同向量的集合。若$S=\{\boldsymbol{0}\}$，则称$S_1,\cdots,S_n$无交连。

• $S=S_1+\cdots+S_n=\{\boldsymbol{x}|\boldsymbol{x}=\boldsymbol{x_1}+\cdots+\boldsymbol{x_n},\boldsymbol{x_i}\in S_i\}$称为$S_1,\cdots,S_n$的和（并）。无交连子空间的和称为子空间的直和，记作$S=S_1\oplus\cdots\oplus S_n$，此时$S$中的向量具有唯一分解$\boldsymbol{x}=\boldsymbol{x_1}+\cdots+\boldsymbol{x_n},\boldsymbol{x_i}\in S_i$。

• 维数公式：$dim(S_1+S_2)=dimS_1+dimS_2-dim(S_1\cap S_2)$。

• 称子空间$S_1,\cdots,S_n$为正交子空间，记作$S_i\perp S_j$，若$\forall\boldsymbol{x_i}\in S_i,\boldsymbol{x_j}\in S_j$，有$\boldsymbol{x_i}\perp\boldsymbol{x_j}$。

•与S正交的所有向量组成向量子空间$S^{\perp}$，称为正交补。$dim(S+S^{\perp})=dim\mathcal{V},\mathcal{V}=S\oplus S^{\perp}$，$\mathcal{V}$中任一向量都可以分解为$S$和$S^{\perp}$的向量和，称为正交分解。

• 若$\forall\boldsymbol{x}\in S$，$\boldsymbol{Ax}\in S$，则称$S$相对于$\boldsymbol{A}$不变。由$\boldsymbol{A}$的特征向量张成的子空间是相对于$\boldsymbol{A}$不变的。零空间$Null(\boldsymbol{A}-\lambda\boldsymbol{I})$是相对于$\boldsymbol{A}$不变的，称之为$\boldsymbol{A}$与特征值$\lambda$对应的特征空间。

• 子空间$\mathcal{H}$和S的夹角$\theta(\mathcal{H},S)=\min\limits_{\boldsymbol{x}\in\mathcal{H},\boldsymbol{y}\in S}\theta(\boldsymbol{x},\boldsymbol{y})$。

• 令$\boldsymbol{P}$是到子空间S的正交投影算子，则对于任一复向量$\boldsymbol{x}$，有$\min\limits_{\boldsymbol{y}\in S}||\boldsymbol{x}-\boldsymbol{y}||_2=||\boldsymbol{x}-\boldsymbol{Px}||_2$，或$\theta(\boldsymbol{x},S)=\theta(\boldsymbol{x},\boldsymbol{Px})$。

• 正交强迫一致问题：$\min\limits_{\boldsymbol{Q}}||\boldsymbol{A}-\boldsymbol{BQ}||_F,s.t.\boldsymbol{Q^TQ}=\boldsymbol{I}$。若有奇异值分解$\boldsymbol{B^TA}=\boldsymbol{U\Sigma V^T}$，则解为$\boldsymbol{Q}=\boldsymbol{UV^T}$，称为$\boldsymbol{B^TA}$的正交极因子。若$\boldsymbol{B}=\boldsymbol{I}$，易知若$\boldsymbol{A}=\boldsymbol{U\Sigma V^T}$，则与$\boldsymbol{A}$最接近的正交矩阵为$\boldsymbol{UV^T}$。

• 矩阵$\boldsymbol{A}$的列空间记为$Col(\boldsymbol{A}) = Span\{\boldsymbol{a_1}, \cdots, \boldsymbol{a_n}\}$，行空间$Row(\boldsymbol{A}) = Col(\boldsymbol{A^H}) = Span\{\boldsymbol{r_1^*}, \cdots, \boldsymbol{r_n^*}\}$。值域$Range(\boldsymbol{A}) = \{\boldsymbol{Ax}\}$，零空间（核）$Null(\boldsymbol{A}) = Ker(\boldsymbol{A}) = \{\boldsymbol{x}|\boldsymbol{Ax} = \boldsymbol{0}\}$，零空间的维数称为$\boldsymbol{A}$的零化维或零度$nullity(\boldsymbol{A}) = dim[Null(\boldsymbol{A})]$。

• $Range(\boldsymbol{A}) = Col(\boldsymbol{A})$。

• $(Row(\boldsymbol{A}))^{\perp} = Null(\boldsymbol{A}), (Col(\boldsymbol{A}))^{\perp} = Null(\boldsymbol{A^H})$。

• $dim(Col(\boldsymbol{A})) = dim(Row(\boldsymbol{A})) = rank(\boldsymbol{A})$，且$rank(\boldsymbol{A}) + dim(Null(\boldsymbol{A})) = n$。

• 初等行变换不改变$Row(\boldsymbol{A})$和$Null(\boldsymbol{A})$，初等列变换不改变$Col(\boldsymbol{A})$和$Null(\boldsymbol{A^H})$。

• 子空间基的构造：初等变换法（P496）、奇异值分解法（P501）

• 两个零空间交的标准正交基：平凡方法、奇异值分解法（P502）

8 投影分析

• 投影定理：令H是向量空间，而M是H内 n 维子空间，若对于H中的向量$\boldsymbol{x}$，在子空间M内有一个向量$\boldsymbol{\hat{x}}$，使得$\boldsymbol{x} - \boldsymbol{\hat{x}}$与$M$中任一向量$\boldsymbol{y}$都满足正交条件，则$||\boldsymbol{x} - \boldsymbol{\hat{x}}|| \leq ||\boldsymbol{x} - \boldsymbol{y}||$对于所有$\boldsymbol{y} \in M$，当且仅当$\boldsymbol{y} = \boldsymbol{\hat{x}}$时成立。

• 向量$\boldsymbol{x}$到子空间M的正交补$M^{\perp}$上的投影则称为正交投影，记为$\boldsymbol{P_M^{\perp}x}$。$\boldsymbol{P_M P_M^{\perp}} = \boldsymbol{P_M P_M^{\perp}} = \boldsymbol{O}$，投影矩阵和正交投影矩阵均奇异。

• 投影的性质（P529）

• 幂 n 矩阵及幂等矩阵性质（P532）

• 考虑$\mathbb{C}^n = S \oplus H$，若$\boldsymbol{x}$有唯一分解$\boldsymbol{x} = \boldsymbol{x_1} + \boldsymbol{x_2}, \boldsymbol{x_1} = \boldsymbol{Px}$，则称$\boldsymbol{P}$是沿着$H$方向到$S$的投影算子，简记为$\boldsymbol{P}_{S|H}$。

• 复向量$\boldsymbol{x}$可以唯一分解为$\boldsymbol{x} = \boldsymbol{Px} + (\boldsymbol{I} - \boldsymbol{P})\boldsymbol{x}$，若$\boldsymbol{P} \perp (\boldsymbol{I} - \boldsymbol{P})$，则称分解为正交分解，$\boldsymbol{P}$为正交投影算子。若不满足$\boldsymbol{P} \perp (\boldsymbol{I} - \boldsymbol{P})$，则称$\boldsymbol{P}$为斜投影算子。

• 线性齐次算子$\boldsymbol{P}$是投影算子，当且仅当$\boldsymbol{P}$是幂等矩阵。

- 齐次线性算子$\boldsymbol{P}$是正交投影算子，当且仅当$\boldsymbol{P}$是是 Hermitian 幂等矩阵。
- 到矩阵$\boldsymbol{A}$的列空间的投影矩阵$\boldsymbol{P}_{\boldsymbol{A}} = \boldsymbol{A}(\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{A})^{+}\boldsymbol{A}^{\boldsymbol{H}}$。