矩阵分析复习提纲

第一章 线性空间与变换

1.1 线性空间

- 定义：加法、乘法闭环+满足8条法则

  交换律，结合律，加0不变，存在相反元素，乘1不变，分配律，...

- 线性相关，线性无关：存在不全为零的数使得两个向量的标准内积为0，为线性相关。
- 线性空间中的基
  - 定义：满足线性无关，其余向量都可以用它表示（可表示性）两个条件；
  - 坐标：就是用基表示的系数组成的向量，有唯一性
  - 坐标变换公式：用不同的基表示同一向量，x是坐标，α, β是基

    $\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} = P \begin{bmatrix} x'_1 \\ x'_2 \\ \dots \\ x'_n \end{bmatrix}$，$(\beta_1, \beta_2, \dots, \beta_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)P$

    P为过渡矩阵。

1.2 线性子空间

- 零空间必为一线性子空间。
- 矩阵的值域与核
  - 值域$R(A) = y|y = Ax, x \in R^n$
  - 核$N(A) = x|Ax = 0$

1.3 线性映射与线性变换

- 定义：满足$\sigma(a_1+s_2)=\sigma(s_1)+\sigma(s_2)$和$\sigma(ks)=k\sigma(s)$
- 线性变换：类似于函数的一种形式，表现为系数的向量矩阵

  例：旋转变换

  $\mathcal{A}(\alpha)=A\alpha, A=\begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}$

  note: A为线性变换的表示矩阵

- 表示矩阵的变换关系：不同基下的相同变换的关系

  $B=P^{-1}AP$，P为两个基的过渡矩阵。

- 线性空间的同构
  - 定义：满射+单射=空间同构

    σ是空间V到U的线性映射，有1) $\sigma(V)=U$; 2) 若$\sigma(\alpha_1)=\sigma(\alpha_2)$，则$\alpha_1=\alpha_2$

第二章 内积空间

2.1 内积运算的定义

1. 实数域

- 记为 (α,β).
- 满足四个条件：交换律、齐次性（就是能提取前面系数）、分配律、非负性。
- 标准内积运算：$(\alpha,\beta)=\alpha^T\beta=\beta^T\alpha=\sum_{i=1}^n a_ib_i$

2. 复数域

- $(\alpha,\beta)=\overline{(\beta,\alpha)}$
- $(k\alpha,\beta)=\overline{k}(\alpha,\beta)$

3. 夹角运算

- $||\alpha||=\sqrt{(\alpha,\alpha)}$为向量的长度

- $<\alpha,\beta>= arccos\frac{(\alpha,\beta)}{||\alpha||||\beta||}, 0\leq <\alpha,\beta>\leq\pi$

2.2 度量矩阵

- n维欧氏空间的基，令$g_{ij}=(\epsilon_i,\epsilon_j), G=(g_{ij})_{n\times n}$，G就为度量矩阵。
- 性质：
    - $A=A^T$，实对称阵。
    - 两个向量在同一组基底下的表示+内积运算：$(\alpha,\beta)=x^TAy$
    - 若是复数，则有：$(\alpha,\beta)=x^HAy$，A是Hermite矩阵$A=A^H$
- 计算：已知一组基底的度量矩阵A和与另一组基底的过渡矩阵C，求另一组基底的度量矩阵B：$B=C^TAC;(B=C^HAC)$

2.3 正交基

- 标准正交基充要条件：度量矩阵是单位矩阵
- Gram-Schmidt正交化方法：

  $\beta_1=\alpha_1$

  $\beta_2=\alpha_2-\frac{(\beta_1,\alpha_2)}{(\beta_1,\beta_1)}\beta_1$

  $\beta_3=\alpha_3-\frac{(\beta_1,\alpha_3)}{(\beta_1,\beta_1)}\beta_1-\frac{(\beta_2,\alpha_3)}{(\beta_2,\beta_2)}\beta_2$
- 单位化：

  $\gamma=\frac{\beta}{||\beta||}$

  note：**在正交化时，注意复数求共轭的运算、出现的比例系数可以丢掉（前提是本向量运算已经完成）**
- 酉矩阵

  由$B=C^HAC$得出，当A，B为单位阵（标准正交基），那么C满足：

  $C^HC=CC^H=I_n$

  C为酉矩阵，为正交矩阵的推广。

2.4 酉变换和Hermite变换

- 定义：不变内积的变换。
- 判断方法：下列命题等价：
    - $\mathcal{A}$是酉变换；
    - $\mathcal{A}(\alpha) = ||\alpha||$；
    - $\epsilon_1, \ldots, \epsilon_n$是标准正交基，那么$\mathcal{A}(\epsilon_1), \ldots, \mathcal{A}(\epsilon_n)$也是标准正交基；
    - $\mathcal{A}$在任意一组正交基下表示的矩阵是酉矩阵。
- Hermite变换：$(\mathcal{A}(\alpha), \beta) = (\alpha, \mathcal{A}(\beta))$
- H-变换判断方法：在标准正交基下的表示矩阵为Hermite矩阵。

2.5 投影

- 投影：$\mathcal{P}(\alpha) = P\alpha$
- 投影矩阵P的判断条件：P为幂等矩阵。
- 投影矩阵P的求解：

  M为S空间的基，N为T空间的基，那么由空间T投影到S空间的投影矩阵的计算为：

  $P = (M, 0)(M, N)^{-1}$

  note：此处的括号不是内积，单纯是矩阵的扩充。
- 正交投影矩阵的充要条件：$P^2 = P = P^H$

第三章 矩阵的Jordan标准型及矩阵分解

Notability上的整理

3.1 Smith标准型

- 矩阵之间的等价条件：A和B经过有限次行和列的初等变换，可以相互转化。等价也称相抵。
- 所有n阶可逆矩阵都可以等价表示为Smith标准型。
- k阶行列式因子$D_k(\lambda)$：n阶可逆矩阵A中所有k阶子式的首项系数为1的最大公因式。
- 不变因子$d_i(\lambda)$：Smith标准型中的每一个对角元素。
- 初等因子：幂指数不为0者

- 相抵判断条件：有相同的初等因子组+秩相等。

3.2 Jordan标准型

- 单纯矩阵：
    - A的那个特征值的几何重数=代数重数，单纯矩阵与对角矩阵相似；
    - 充要条件是所有初等因子都是1次的。
- 相似$A \sim B$充要条件：有相同的不变因子/初等因子组。

3.3 相似对角化变换矩阵

- $P^{-1}AP = J$中求P用$AP = PJ$
- **Note:需要保证非齐次方程有解，需满足rank(左式)=rank[左式，右列]**

3.4 满秩分解

- 总结：选取行简化阶梯型矩阵主元所在的列对应的列向量构成满秩矩阵，将阶梯型矩阵全为零的行去掉后即可构成行满秩矩阵。
- 满秩分解不唯一：但乘积保持不变，广义逆矩阵

  $G^H(GG^H)^{-1}(FF^H)^{-1}F^H$

3.5 谱分解

- 正规矩阵
    - 定义：满足$A^TA = AA^T(A^HA = AA^H)$
    - 结构定理，正规矩阵的充要条件：存在U使得$U^HAU = diag(\lambda_1, \ldots, \lambda_n)$
- 正规矩阵和单纯矩阵谱分解的区别：单纯矩阵求出特征向量后不用正交单位化。

3.6 奇异值分解

- 奇异值定义：$\sigma_i = \sqrt{\lambda_i(A^HA)}$
- 奇异值分解形式：$A \in C_r^{m\times n}$

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

  $\Sigma = diag(\sigma_1, ..., \sigma_r)$，都是正奇异值。

  U，V都是酉矩阵。

3.7 最小多项式 Cayley-Hamilton定理

- CH定理：$\phi(\lambda) = det(\lambda I - A)$，则$\phi(A) = 0$
- 化简矩阵多项式：先求$\phi(\lambda) = det(\lambda I - A)$的表达式，之后对应题目中的$f(\lambda)$进行化简
- 最小多项式：A的化零多项式中次数最低的首1多项式，记为$m_A(\lambda)$
- 最小多项式求法：化为Jordan标准型，取分块矩阵最小多项式的最小公倍式

第四章 范数理论

4.1 向量范数

- 定义（向量范数三公理）：
  - 正定性：当$x \neq 0$时，有$||x|| > 0$;
  - 齐次性：$||kx|| = |k|||x||$
  - 三角不等式：$||x+y|| \leq ||x|| + ||y||$
- 经典常用范数
  - 1-范数：$||x||_1 = \sum_{k=1}^n |x_k|$
  - 2-范数：$||x||_2 = \sqrt{\sum_{k=1}^n |x+k|^2} = \sqrt{x^Hx}$
  - ∞-范数：$||x||_\infty = max_k|x_k|$
  - p-范数：$||x||_p = (\sum_{k=1}^n |x_k|^p)^{\frac{1}{p}}$
- 不等式（乘积的和小于和的乘积）
  - 【乘积】Holder不等式：
    $\sum_{k=1}^n |x_k||y_k| \leq (\sum_{k=1}^n |x_k|^p)^{\frac{1}{p}}(\sum_{k=1}^n |y_k|^p)^{\frac{1}{p}}$,
    $p > 1, q > 1, \frac{1}{p} + \frac{1}{q} = 1$

- Cauchy-Schwarz不等式：
  $\sum_{k=1}^{n}|x_k||y_k| \le (\sum_{k=1}^{n}|x_k|^2)^{\frac{1}{2}}(\sum_{k=1}^{n}|y_k|^2)^{\frac{1}{2}}$
- 【和】Minkowski不等式：
  $(\sum_{k=1}^{n}|x_k+y_k|^p)^{\frac{1}{p}} \le (\sum_{k=1}^{n}|x_k|^p)^{\frac{1}{p}} + (\sum_{k=1}^{n}|y_k|^p)^{\frac{1}{p}}, p \ge 1$

- 范数的等价性
  - 存在正数c使得

    $c_1||x||_b \le ||x||_a \le c_2||x||_b$

    则称a b两范数等价

  - n维线性空间中任意两个向量范数都是等价的

- 向量的收敛性
  - 向量收敛：所有元素都收敛
  - 证明：此向量应用任意范数，都有$lim_{k\to\infty}||x^{(k)}-x||=0$

4.2 矩阵范数

- 定义：多了相容性

  $||AB|| \le ||A||||B||$

- 范数种类
  - m1范数：$||A||_{m_1} = \sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|$

  - Frobenius范数：
    $||A||_F = ||A||_{m_2} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2} = \sqrt{tr(A^HA)}$

    note: 矩阵的迹等于矩阵的特征值之和，故F范数就是矩阵的奇异值之和

  - ∞-范数：$||A||_{m_\infty} = max(m,n)max_{i,j}|a_{ij}|$

- 与向量范数的交叉融合
  - 相容定义：同上，就是把其中一个矩阵范数变成向量范数
  - 矩阵范数必存在与它相容的向量范数

4.3 算子范数

- 由来：为了找出一种矩阵范数，有$||I_n||=1$

- 定义：$||*||_v$是向量范数，$\forall A \in C^{n\times n}$,

$||A|| = max_{x \neq 0} \frac{||Ax||_v}{||x||_v}$

- 范数种类
    - 1-范数，列模和范数：$||A||_1 = max_j \sum_{i=1}^{n} |a_{ij}|$
    - 2-范数，谱范数：$||A||_2 = \sqrt{max_i \lambda_i (A^H A)}$，最大正奇异值
    - ∞-范数，行模和：$||A||_\infty = max_i \sum_{j=1}^{n} |a_{ij}|$
- 2-范数的特殊性质（算子范数）
    - $||A^H||_2 = ||A||_2$
    - 【酉不变性】U,V为酉矩阵，
    - A为正规矩阵，则$||A||_2 = max_j |\lambda_j|$，A的谱半径

4.4 谱半径

- 定义：A是方阵，$\rho(A) = max_j |\lambda_j|$
- 性质
    - $\rho(A^k) = [\rho(A)]^k$
    - $\rho(A^H A) = \rho(AA^H) = ||A||_2^2$
    - A为正规矩阵时，$\rho(A) = ||A||_2$
    - 对任一矩阵范数，都有$\rho(A) \leq ||A||$

第五章 矩阵分析

5.1 矩阵序列

- 符号：$A^{(k)}$表示矩阵序列A中第k和矩阵
- 收敛：所有元素都收敛，充要条件是任一矩阵范数都收敛
- 收敛矩阵
    - 定义：A为方阵，若$lim_{k \to +\infty} A^k = 0$，A为收敛矩阵
    - 充要条件：$\rho(A) < 1$
    - 推论：若存在一种范数，使$||A|| < 1$，则A为收敛矩阵。

5.2 矩阵级数

- 定义：无穷和$A^{(0)} + A^{(1)} + ... + A^{(k)}$，记为$\sum_{k=0}^{+\infty} A^{(k)}$
- 收敛性质：
    - $lim_{k \to +\infty} A^{(k)} = 0$
    - 可拆分
- 绝对收敛
    - 定义：所有项的所有元素都收敛
    - 判别条件：$\sum_{k=0}^{+\infty} ||A^{(k)}||$收敛
- 幂级数
    - 定义：$\sum_{k=0}^{+\infty} c_k A^k = c_0 I + c_1 A + c_2 A^2 + ... + c_k A^k + ...$
    - 敛散性判定：
        - 收敛半径R：$\frac{1}{R} = lim_{k \to +\infty} |\frac{c_{k+1}}{c_k}|$
        - 当$\rho(A) < R$，绝对收敛
        - 当$\rho(A) > R$，发散
    - Neumann级数
        - 定义：幂级数的系数为1，$\sum_{k=0}^{+\infty} A^k$
        - 收敛条件：$\rho(A) < 1$
        - 和：$\sum_{k=0}^{+\infty} A^k = (I - A)^{-1}$

5.3 矩阵函数

- 定义：$f(A)$，正常函数怎么套用就怎么用
- Jordan标准型法求函数值
    - A是对角阵/相似于对角阵：转换后，把函数全作用在对角元素上
    - A相似于Jordan标准型：作用在Jordan块上

      上三角元素：$\frac{1}{(n_i-1)!} f^{(n_i-1)}(\lambda_i)$

      note：当函数带有t时，要注意自变量是矩阵A，而不是t，要将t看作系数去求导
- 矩阵多项式法求函数值（有限级数法、待定系数法）
    - 求特征值，个数为n
    - 列余式：$r(\lambda) = a_0 + a_1 \lambda + ... + a_n \lambda^{n-1} = f(\lambda t)$
    - 求导数，用λ表示$a_0, ..., a_n$

    - 代入原式：$f(At)=a_0I+a_1A+...+a_nA^n$
- 单纯矩阵的谱分解法求函数值
    - 谱分解结束后，直接将函数作用到特征值上，再相加

5.4 矩阵微分

- 函数矩阵，矩阵值函数
    - 定义：所有元素都是定义在某一区间的实函数
    - 可逆、可导、可积都取决于每个元素，运算法则也相同

      note：矩阵乘法的顺序

      若n阶可逆函数矩阵在[a,b]上可导，则 $\frac{d}{dt}A^{-1}(t)=-A^{-1}(t)(\frac{d}{dt}A(t))A^{-1}(t)$