特征值和特征向量

引子

- 线性映射中前后共线的向量 $Ax = \lambda x \Rightarrow (\lambda I - A)x = 0$有非零解
  - $\det(\lambda I - A) = 0$
- 代数学基本定理及其推论
  - 代数学基本定理
    - 复系数一元n次多项式$\mathbb{C}$上至少有一个根
    - 推论
      - 复系数一元n次多项式$p(x)$，在$\mathbb{C}$上恰好有n个根（可能相同），即存在因式分解$p(x) = a_0(x - x_1)^{n_1}(x - x_2)^{n_2}\cdots(x - x_s)^{n_s}$，其中$n_1 + n_2 + \cdots + n_s = n$，$n_i$称为重数，$x_i$称为$p(x)$的$n_i$重根
  - 韦达定理
    - 复系数一元n次多项式$p(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n$的根（计重数），$x_1, x_2, \cdots, x_n$满足
      - $-\dfrac{a_1}{a_0} = x_1 + x_2 + \cdots + x_n$
      - $(-1)^k\dfrac{a_k}{a_0} \quad \sum_{1 \le i_1 < \cdots < i_k} x_{i_1}x_{i_2}\cdots x_{i_k}$
      - $\vdots$
      - $(-1)^n\dfrac{a_n}{a_0} = x_1x_2\cdots x_n$

基本概念

- 特征值
  - 给定n阶方阵A，如果对$\lambda \in \mathbb{C}$，存在非零向量$x \in \mathbb{C}^n$，使得$Ax = \lambda x$
    则称数λ为方阵A（在$\mathbb{C}$上）的一个特征值，而称非零向量x为方阵A（在$\mathbb{C}$上）的一个属于特征值λ的特征向量
    二元组（λ, x）常称为方阵A的一个特征对
    - 注意：
      1.只有方阵才有特征值和特征向量
      2.零向量不是特征向量
      3.如果x是特征向量，则对任意$k \neq 0$，kx都是特征向量
- 盖尔（Gershgorin）圆盘定理
  - 对矩阵$A = [a_{ij}]_{n\times n}$，定义如果n个圆盘
    $G_i(A) = \{z \mid |z_{ii}| \le \sum_{j\neq i}|a_{ij}|\}$
    那么矩阵的任意特征值λ一定落在某个圆盘中. 因此，矩阵的全部特征值一定落在这n个圆盘中
- 特征值的多项式特征
  - 代数重数
    - 给定n阶方阵A及A的一个特征值$\lambda \in \mathbb{C}$，如果λ_0是特征多项式$p_A(\lambda)$的n_0重根，
      则称n_0为λ_0作为A的特征值的代数重数（简称重数），称λ_0的一个n_0重特征值
      一个1重特征值，又称为单特征值
      - 给定n阶实方阵A，如果λ_0是它的一个非实数特征值，则$\overline{\lambda_0}$也是它的特征值，
        且其代数重数和λ_0的代数重数相等
        进一步地，如果复向量x_0是属于λ_0的特征向量，则$\overline{x_0}$是属于$\overline{\lambda_0}$的特征向量
  - 迹，行列式与特征值的关系
    - 迹
      - 给定n阶方阵A，其特征多项式M有如下形式：
        $p_A(\lambda) = \lambda^n - trace(A)\lambda^{n-1} + \cdots + (-1)^n\det(A)$
        其中$trace(A) = a_{11} + a_{22} + \cdots + a_{nn}$是方阵$A$的对角元素的和，称为方阵$A$的迹
    - 迹，行列式与特征值的关系
      - 由韦达定理得：
        $\lambda_1 + \lambda_2 + \cdots + \lambda_n = trace(A)$
        $\lambda_1\lambda_2\cdots\lambda_n = \det(A)$
  - 弗朗西斯（Francis）算法和多项式根值逼近
    - Francis算法
      - 实际生活中求解特征值问题较为困难，往往要使用Francis算法求解特征值
    - 多项式根值逼近
      - $p(x) = c_0 + c_1x + \cdots + c_{n-1}x_{n-1} + x_n$，令
        $y = \begin{bmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{n-1} \end{bmatrix}, C_p = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ -c_0 & \cdots & -c_{n-2} & -c_{n-1} \end{bmatrix}$
        容易验证$C_py = \lambda y$当且仅当$p(\lambda) = 0$
        即λ是多项式$p(x)$的根，当且仅当λ是矩阵C_p的特征值
        - 如此，C_p的特征多项式是$p(x)$，估计$p(x)$的根转化为了求C_p的特征值，C_p又称作$p(x)$的友矩阵

对角化和谱分解

- 定义
  - 对方阵A，如果存在可逆矩阵X使得$X^{-1}AX = \Lambda$是对角矩阵，则称A是（在C）上可对角化的，称把A对角化或X对角化A
    若方阵A，X都是实矩阵，则称A在$\mathbb{R}$上可对角化
    当A可对角化时，分解$A = X\Lambda X^{-1}$称为A的谱分解
- 几何重数
  - 定义
    - 给定n阶方阵A及其特征值λ_0，称特征子空间$\mathcal{N}(\lambda_0I_n - A)$的维数为$\lambda_0$作为特征值的几何重数
  - 半单和亏损特征值
    - 几何重数和代数重数相等的特征值，称为半单特征值
    - 几何重数小于代数重数的特征值，称为亏损特征值
- 条件
  - 充要条件
    - 对n阶方阵A，A可对角化，当且仅当A有n个线性无关的特征向量
      - 何时特征向量线性无关？
        - 方阵属于不同特征值的特征向量线性无关
          - 充分条件
            - 有n个不同特征值的n阶矩阵，即特征值都是单特征值的方阵，可对角化
        - 当方阵A的特征值的代数重数大于1时
          - 引入几何重数判别特征向量到底是否线性无关？
            - n阶方阵A可对角化，当且仅当其特征值都半单

相似

- 定义
  - 对方阵A, B，如果存在可逆矩阵X使得$X^{-1}AX = B$，则称A和B相似，或A相似于B
- 本质
  - 相似的本质是线性变换在不同基下的效果等价
- 相似不变量
  - 1.秩
  - 2.特征多项式、特征值、特征值的代数重数、迹、行列式
  - 3.特征值的几何重数
- Jordan分解
- 同时对角化

FxMarkBrown's Blog