第一章：线性空间

1.1 线性空间

数域 (4条)

满足对加减乘除**四则运算封闭**

常见的数域：

- $\mathbb{Q}$ 有理数
- $\mathbb{R}$ 实数
- $\mathbb{C}$ 复数

自然数$\mathbb{N}$和整数$\mathbb{Z}$不是数域

向量空间 (2条)

- 对加法封闭：$\boldsymbol{\alpha}+\boldsymbol{\beta}\in V$
- 对数乘封闭：$k\boldsymbol{\alpha}\in V$

加群（4条）

- 交换律：$a+b=b+a$
- 结合律：$a+(b+c)=(a+b)+c$
- 存在零元素$\boldsymbol{\theta}$：$\boldsymbol{\alpha}+\boldsymbol{\theta}=\boldsymbol{\alpha}$
- 存在负元素：$\boldsymbol{\alpha}+(-\boldsymbol{\alpha})=\boldsymbol{\theta}$

线性空间（8条）

在**加群**的基础上满足额外的四条数乘性质：

- 数乘分配律： $\lambda(\boldsymbol{\alpha}+\boldsymbol{\beta})=\lambda\boldsymbol{\alpha}+\lambda\boldsymbol{\beta}$
- 数乘分配律：$(\lambda+\mu)\boldsymbol{\alpha}=\lambda\boldsymbol{\alpha}+\mu\boldsymbol{\alpha}$
- 数乘结合律：$(\lambda\mu)\boldsymbol{\alpha}=\lambda(\mu\boldsymbol{\alpha})$
- 数乘单位元：$1\cdot\boldsymbol{\alpha}=\boldsymbol{\alpha}$

上面四条是对线性空间对应的数域的约束，而不是对线性空间中元素的约束，对于常规定义的数域，上述几条自动满足，一个线性空间中不需要单位元素。

对常规线性空间，只需验证**加法封闭、含零向量、数乘封闭**这三条性质即可。

常见的线性空间

- 向量空间
- 矩阵空间
- 一元多项式集合：$P_n(x)=\{\sum a_i x^i|a_i\in\mathbb{C}\}$
- $\mathbf{Ax}=\mathbf{b}$的解集
- $[a,b]$上的全体多项式S_1,全体可微函数S_2,全体连续函数S_3,全体可积函数S_4,全体实函数S_5

- $S_1 \subset S_2 \subset S_3 \subset S_4 \subset S_5$

线性空间的性质

- 零元素**唯一**
- 每个元素对应的负元素**唯一**
- $k\boldsymbol{\alpha} = \boldsymbol{\theta}$只有两种可能性，要么元素为0，要么$k$为零元素
- 加法和数乘合称为**线性运算**

1.2 线性子空间

子空间(2条)

- 子空间本身也需要是一个空间
- **子空间必须要有零元素，满足加法与数乘封闭**
- 包含空间本身和单独一个零元素的空间也为子空间，称之为平凡子空间

子空间判别法

下面三个命题等价

- W为V的子空间
- $k\boldsymbol{\alpha} \in W$ & $\boldsymbol{\alpha} + \boldsymbol{\beta} \in W$
- $k\boldsymbol{\alpha} + l\boldsymbol{\beta} \in W$

和空间和交空间

- 交空间 $W_1 \cap W_2$ 同时包含W_1与W_2中元素的最大子空间【bool操作减法】
- 和空间 $W_1 + W_2$ 同时包含W_1与W_2的最小子空间 【升维加法操作】
- 有限个线性子空间的**和**与**交**仍是线性子空间

张成子空间

由多个向量张成的子空间：$\operatorname{span}(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \cdots, \boldsymbol{\alpha_n})$

相当于：$W = \{k_1\boldsymbol{\alpha_1} + \cdots k_n\boldsymbol{\alpha_n} | k_i \in F\}$

矩阵零空间列空间

- 零空间/核空间：$N(\mathbf{A}) = \{\mathbf{x} \in \mathbb{C}^n | \mathbf{Ax} = \mathbf{0}\}$
    - 相当于方程组的解集
- 列空间/值空间：$R(\mathbf{A}) = \{\mathbf{y} \in \mathbb{C}^m | \mathbf{y} = \mathbf{Ax}\}$
    - 若$\mathbf{A} = [\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \cdots, \boldsymbol{\alpha_n}]$ ，那么 $R(\mathbf{A}) = \operatorname{span}(\boldsymbol{\alpha_1}, \boldsymbol{\alpha_2}, \cdots, \boldsymbol{\alpha_n})$

1.3 基与坐标

线性相关&无关

线性相关需要满足 $\sum k_i\boldsymbol{\alpha}_i=\boldsymbol{\theta}$ 只有在$k_i=0$时成立。

单个零向量线性相关，单个非零向量线性无关

极大线性无关组与秩

秩：$\text{rank}[\boldsymbol{\alpha}_i]=r$

即对于$\boldsymbol{\alpha}_i$中的所有元素，都可以被r个$\boldsymbol{\alpha}_{i_j}$表示，$\boldsymbol{\alpha}_{i_j}$即为**极大线性无关组**

基(2条)

- 构成基的向量组$\boldsymbol{\alpha}_i$线性无关
- 线性空间V中的任意向量都可以被这个基描述

基一定是极大线性无关组，但是极大线性无关组不一定是基

唯一表示定理：线性空间V中的任意向量都可以有其任意的基唯一表示

过渡矩阵

- 对于一个线性空间上的两组基$\mathbf{y}_i$和$\mathbf{x}_i$: $[\mathbf{y}_i]_{1\times n}=[\mathbf{x}_i]_{1\times n}\mathbf{A}$，$\mathbf{A}$为过渡矩阵
- 过渡矩阵可逆

维数

线性空间V中，不同线性无关组中向量个数最大者叫做V的维数，记为$\dim V$

可直接用线性空间中基中包含的向量个数定义该空间的维数

基扩充定理

n线性空间中，任意n个线性无关的向量均构成了一个基底。

维数定理

$\dim(W_1+W_2)=\dim(W_1)+\dim(W_2)-\dim(W_1\cap W_2)$

1.4 内积空间

内积空间/酉空间（4条）

某个线性空间上定义了内积，并满足下面四条，根据数域是$\mathbb{R}$或$\mathbb{C}$，称其为内积或者酉空间

- 共轭性：$(x,y)=\overline{(y,x)}$ 共轭对称
- 可加性：$(x+y,z)=(x,z)+(y,z)$
- 齐次性：$(kx,y)=k(x,y)$
- 正定性：$(x,x)\geq 0$ 且仅当$x=0$时内积为零

注意：$(x,ky)=\overline{(ky,x)}=\bar{k}(x,y)$

几个例子：

在$\mathbb{R}^n$中：$(x, y) = y^T x$

在$\mathbb{C}^n$中：$(x, y) = y^H x$

在$\mathbb{C}^n$中：$(x, y) = y^H A x$，此时矩阵需要为Hermite矩阵$A^H = A$，并正定$x^H A x \geq 0$，仅当$x = 0$时$x^H A x = 0$

Hermite矩阵

- 对于向量$z = a + bi$，取共轭$z^H = a - bi$
- Hermite矩阵：满足$A^H = A$，取共轭后同时求转置
- 反Hermite矩阵：满足$A^H = -A$
- **共轭转置的计算**：$(AB)^H = B^H A^H$

正定矩阵

- $\mathbf{x} = [x_1, \cdots, x_n]^T \in \mathbb{C}^n$ ，$f(\mathbf{x}) = \mathbf{x}^H A \mathbf{x}$
- 矩阵A为正定矩阵&$f(\mathbf{x})$为正定二次型：对于任意的$\mathbf{x}$，仅有$\mathbf{x} = 0$时$f(\mathbf{x}) = 0$，其与都$f(\mathbf{x}) \geq 0$
- 矩阵A为半正定矩阵&$f(\mathbf{x})$为半正定二次型：对于任意的$\mathbf{x}$，$f(\mathbf{x}) \geq 0$
- 顺序主子式的行列式都大于(等于)零
- 对于实对称矩阵/Hermite矩阵，若所有特征值大于零，则为正定矩阵
- Hermite矩阵的特征值都必定为实数

度量矩阵（Gram矩阵）

ϵ_i为内积空间的一组基，那么下面的矩阵A即为度量矩阵

$$A = \begin{bmatrix} (\epsilon_1, \epsilon_1) & (\epsilon_1, \epsilon_2) & \cdots & (\epsilon_1, \epsilon_n) \\ \vdots & \vdots & \vdots & \vdots \\ (\epsilon_n, \epsilon_1) & (\epsilon_n, \epsilon_2) & \cdots & (\epsilon_n, \epsilon_n) \end{bmatrix}$$

若向量$\mathbf{x}, \mathbf{y}$在同一组基ϵ_j下的投影坐标构成的向量为ξ和η，那么$\mathbf{x}, \mathbf{y}$的内积可表示为：

$$(\mathbf{x}, \mathbf{y}) = \eta^H A^H \xi$$

度量矩阵的性质：

- Hermite矩阵
- 正定矩阵
- 同一内积空间内不同基下的度量矩阵**相合**：$P^H G(\epsilon_j) P = G(\xi_k)$

（广义）长度

- 这个长度的定义可以延拓至一个高维空间，假设对于一个四维空间，可能无法几何上面想象，可以将其当作一个状态空间，那么存在这个四维向量，其广义的长度可以理解为从其中一个状态到另外一个状态的某种度量
- $\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})}$
- 满足正定性，齐次性$\|k\mathbf{x}\| = |k| \, \|\mathbf{x}\|$
- 满足平行四边形法则：$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)$

- 满足三角不等式：$\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

柯西(Cauchy-Schwarz)不等式

- $|(\mathbf{x},\mathbf{y})| \leq \|\mathbf{x}\| \|\mathbf{y}\|$
- 等于的情况仅在$\mathbf{x},\mathbf{y}$线性相关时成立

向量夹角

$$\alpha = \cos^{-1}\frac{(\mathbf{x},\mathbf{y})}{\|\mathbf{x}\| \|\mathbf{y}\|} \in [0,\pi]$$

正交和正交向量组

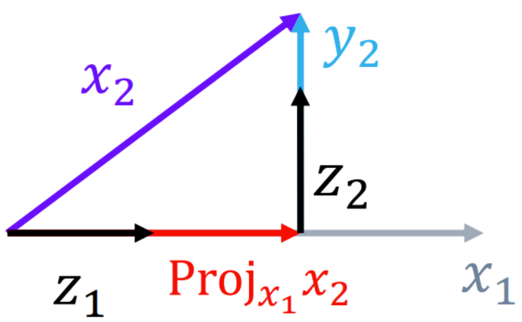
- 即满足$(\mathbf{x},\mathbf{y})=0$
- 标准即两向量为单位向量
- **零向量与任何向量均正交，正交向量组需要排除这个作弊的零向量**
- 当且仅当两向量正交时，$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ 成立，这即为广义的勾股定理

正交向量组线性无关

- 即存在$\exists k_i \neq 0, \sum k_i\mathbf{x}_i = \boldsymbol{\theta}$
- 在n维内积空间中，正交向量组中的向量不会超过n个

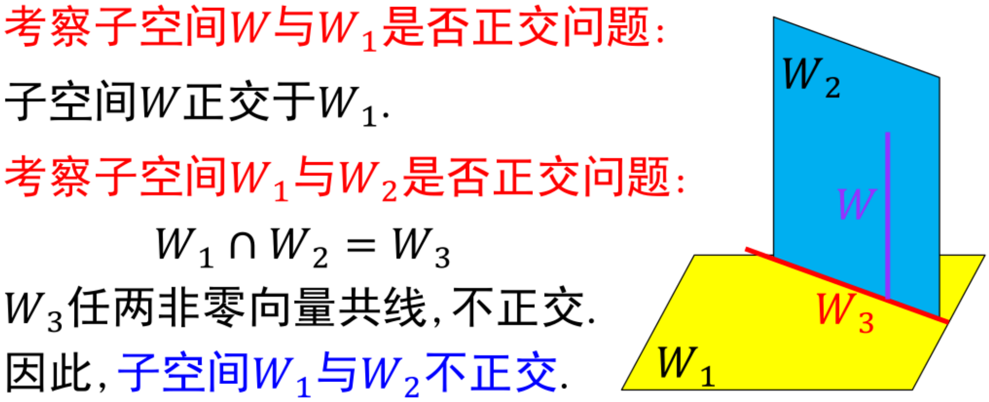
正交基和标准正交基

- n维内积空间中，n个(单位)正交向量组构成(标准)正交基
- 在两个**标准正交基**下，两个向量的坐标分别为$\boldsymbol{\alpha},\boldsymbol{\beta}$，那么：$(\mathbf{x},\mathbf{y})=\boldsymbol{\beta}^H\boldsymbol{\alpha}$
- 有限维内积空间必存在标准正交基，可以通过Gram-Schmidt方法构建
- Gram-Schmidt方法：
    - 取第一个基底$\mathbf{x}_0$
    - 计算后续$n-1$个基底：$\mathbf{y_k} = \mathbf{x_k} - \sum_{i=1}^{k-1}\frac{(\mathbf{x}_k,\mathbf{y}_i)}{(\mathbf{y}_i,\mathbf{y}_i)}\mathbf{y}_i$



向量集合的布尔运算

- 向量正交于集合：$\mathbf{x} \perp W$，其中W为内积空间V的子集
- 两集合正交：$W_1 \perp W_2$，即两个集合中的任意向量正交



- 子空间正交补：$W^{\perp} = \{\mathbf{x} \in V | \mathbf{x} \perp W\}$
  - $V = W + W^{\perp}$
  - W和$W^{\perp}$都是线性子空间

1.5 直和与投影

直和与正交直和

- 空间$W_1 + W_2$的任意向量唯一地表示为W_1与W_2的向量之和，则$W_1 + W_2$为W_1与W_2的直和，记作$W_1 \dot{+} W_2$
- 特别地，$V = W_1 \dot{+} W_2$称为直和分解
- 若同时$W_1 \perp W_2$，则称为正交直和，记作$W_1 \oplus W_2$

直和判定定理

四个命题等价：

- $W_1 + W_2$为直和
- $W_1 + W_2$中的零元素表达法唯一
- $W_1 \cap W_2 = \{\boldsymbol{\theta}\}$
- $\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2)$

直和的性质和定理

在$A^2 = A$情况下：

- $R(A) \dot{+} R(I - A) = \mathbb{C}^n$
- $\operatorname{rank}(I - A) = n - r$

性质和定理

- $W_1 \perp W_2$，则 $W_1 + W_2 = W_1 \oplus W_2$
- $V = W \oplus W^{\perp}$
- 对于任意方阵$A \in \mathbb{C}^{n \times n}$，$N(A) \oplus R(A^H) = \mathbb{C}^n$

- $(U+W)^{\perp} = U^{\perp} \cap W^{\perp}$

秩-零化定理 & 亏加秩定理

$$\dim N(A) + \dim R(A) = \dim(V) = n$$

$$\dim N(A) + \operatorname{rank} A = \dim(V) = n$$

另一个角度理解，在**向量层面**对于$Ax = 0$：

$$解空间的维数\dim N(A) = 矩阵A的列数n - 矩阵A的秩\operatorname{rank}(A)$$

记忆时：$r = \operatorname{rank} A = \dim R(A)$

投影与正交投影

- $V = W_1 \dot{+} W_2$，$z = x + y$，$x \in W_1, y \in W_2$，此时x为W_1上的投影
- 若$V = W_1 \oplus W_2$，此时x为W_1上的正交投影
- 若$W = \operatorname{span}(\mathbf{x_i})$，那么$\operatorname{Proj}_W \mathbf{a} = \sum \frac{(\mathbf{a}, \mathbf{x_i})}{\|\mathbf{x_i}\|^2} \mathbf{x_i}$

第二章：线性映射与矩阵

2.1 映射与多项式

基本映射概念

- $y = f(x)$：从$x \in V$定义域映射到$y \in W$值域，y为x的像，x为y的原像
- 单射、满射、双射（一一映射）
    - 单射必须对应的Y不同，满射必须Y被射满

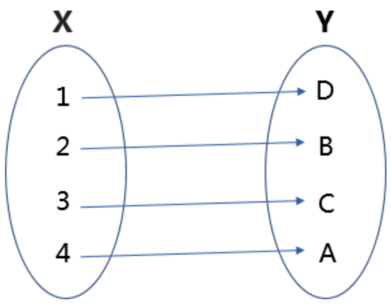


图1双射（单射与满射）

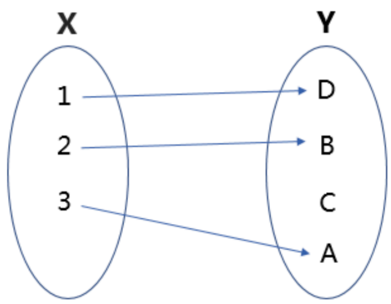


图2 单射但非满射

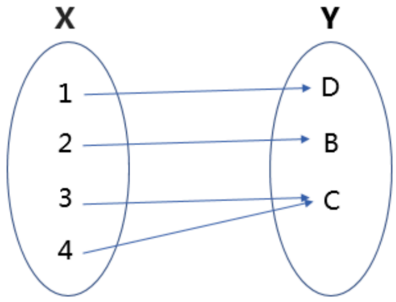


图3 满射但非单射

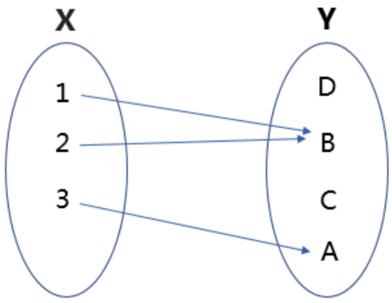


图4 非满射非单射

- 映射相等（$|x|, \sqrt{x^2}$）
- 映射乘积（$f_2(f_1(x))$）
- 可逆映射（反函数）
  - 可逆映射唯一
  - 充分必要条件为双射
- 变换：到自身的映射
- 一一变换/置换：到自身的双射

多项式

- 一元多项式：$g(\lambda) = a_n\lambda^n + \cdots + a_1\lambda + a_0$
- $g(\lambda)$和$h(\lambda)$存在不为零，那么存在唯一的多项式$p(\lambda), q(\lambda)$，满足：
  - $g(\lambda) = p(\lambda)h(\lambda) + q(\lambda)$
  - $q(\lambda)$要么为零，要么其阶数小于$h(\lambda)$
  - 为零时即多项式整除，记为$h(\lambda)|g(\lambda)$
- 公因式、公倍式
- 友矩阵

2.2 线性映射

线性映射

线性映射满足可加性与齐次性：

- $T(x + y) = T(x) + T(y)$
- $T(\lambda x) = \lambda T(x)$

常见的线性映射：

- 零变换$T(x) = \theta$
- 恒等变换$T(x) = x$
- 负变换$T(x) = -x$
- 伸缩、反射、旋转
- 微分变换、积分变换、正交投影变换

- $T(A)=A^T, T(A)=AB; A\in\mathbb{C}^{n\times n}$

注：给定两个线性变换，$T_2T_1=T_2(T_1(x))$

线性映射的性质

定理1：从可加性和齐次性可以推出：

$$T(\sum k_i\mathbf{x}_i)=\sum k_iT(\mathbf{x}_i)$$

- $T(\theta)=\theta', \theta\in V, \theta'\in W$ **即线性映射必须保持原点不动，平移变换不是线性变换**
- $T(-x)=-T(x)$
- 变换前后的向量若线性相关/线性无关，变换后的向量**依旧**线 性相关/线性无关

定理2：仅当T是单射时,定义域V中线性无关向量组映射到W时依旧是线性无关的

假设一个线性映射做升维操作：$(x,y)\rightarrow(x,y,0)$，此时满足单射，两个向量(x,y)假设线性无关不共线，映射后的两向量$(x,y,0)$虽然共面，但是依旧不共线，是线性无关的。

若线性映射做升维操作，但是不是单射$(x,y)\rightarrow(x,0,0)$，此时可以发现被映射的向量肯定在一个轴上，原先线性无关的向量变为线性相关。

这也意味着：**线性映射不一定将一组基映射为像空间的一组基**

线性映射空间

集合$L(V,W)$赋以加法和数乘构成数域F上的线性空间，称为线性映射空间，$L(V)$构成线性变换空间

$$T+S\in L(V,W)$$

$$\lambda T\in L(V,W)$$

且满足8条线性空间运算规则

2.3 矩阵与同构

线性映射与矩阵的联系

- ϵ_i为V的基，存在$T\in L(V,W)$，$x=\sum x_i\epsilon_i$，那么：

$$T(x)=\sum x_iT(\epsilon_i)$$

- ϵ_i,η_i分别为V,W的基，存在$T\in L(V,W)$，那么每组η_i都可以被$T(\epsilon_i)$线性表示：

$$\begin{gathered}T(\epsilon_1)=a_{11}\eta_1+a_{21}\eta_2+\cdots+a_{m1}\eta_m\\ T(\epsilon_2)=a_{12}\eta_1+a_{22}\eta_2+\cdots+a_{m2}\eta_m\\ \vdots\\ T(\epsilon_n)=a_{1n}\eta_1+a_{2n}\eta_2+\cdots+a_{mn}\eta_m\end{gathered}$$

- 这样即定义了线性映射$T\in L(V,W)$在基ϵ_i和η_i下的矩阵：

$$[T(\epsilon_1,\cdots,\epsilon_n)]=[\eta_1,\cdots,\eta_m]A$$

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in F^{m\times n}$$

注意，这里的矩阵和前面的方程系数的对应关系，有个转置关系

定理1：在两个空间的基都给定的条件下，线性映射T唯一决定了A

同构映射

同构映射将两个本质完全相同的集合进行映射。

比方说，将1×3的数组集合$\{(a,b,c)\}$映射到三维笛卡尔空间，每个元素代表空间上的一个点。

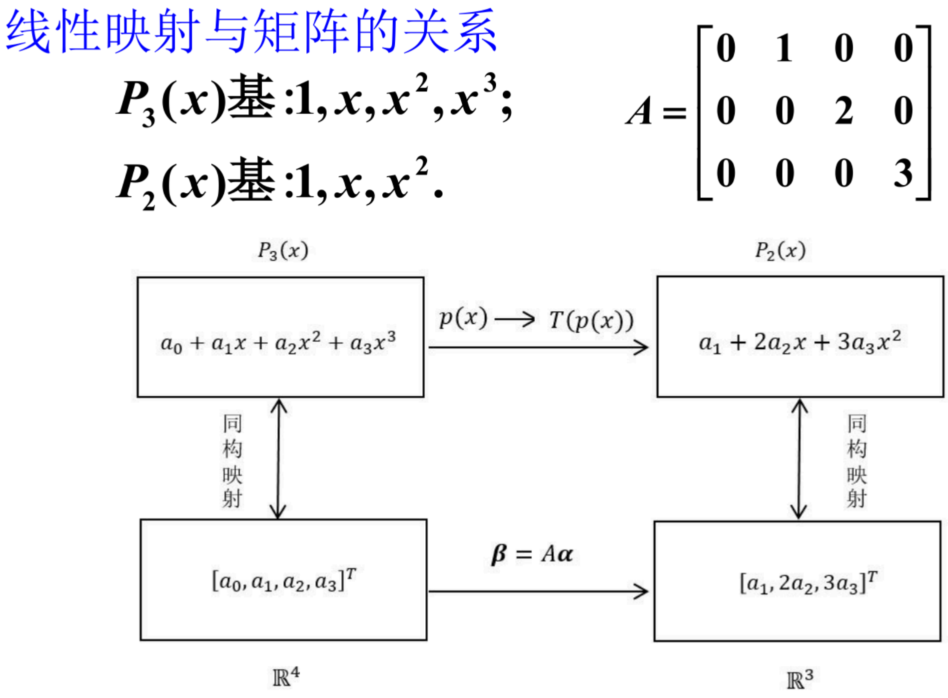
- V,W为数域F上的线性空间，存在双射$f:V\to W$满足：
  - $f(x+y)=f(x)+f(y)$
  - $f(\lambda x)=\lambda f(x)$
- 则f是V到W的同构映射，并称线性空间V,W同构
- **定理2**：线性映射空间$L(V,W),\dim(V)=n,\dim(W)=m$和矩阵空间$F^{m\times n}$同构

同构映射具有以下的性质：

- 零元素相同：$T(\theta)=\theta'$
- $T(-x)=-T(x)$
- $T(\sum a_ix_i)=\sum a_iT(x_i),a_i\in F,x_i\in V$
- 由于像和原像本质是相同的，V的向量组仅在其像线性相关时，线性相关
- 同理，ϵ_i为V的一组基，那么$T(\epsilon_i)$同样也是W的一组基
- T^{-1}存在且同样时同构映射

定理3：线性空间同构当且仅当它们的维数相等

定理4：$x=[\epsilon_1\cdots\epsilon_n]\alpha$，$T(x)=[\eta_1\cdots\eta_m]\beta$，那么有：$\beta=A\alpha$



定理5：线性映射在不同基下的矩阵是相抵的，记为$A\cong B$

- 若$\epsilon_i'=\epsilon Q,\eta_i'=\eta_i P$，
- 且$T(\epsilon_i)=\eta_i A,T(\epsilon_i')=\eta_i' B$
- 那么：$B=P^{-1}AQ$，即矩阵A,B相抵
- 相似变换反映的是同一个线性变换

定理6：由于矩阵和线性变换是同构的，亏加秩定理可以同时在矩阵A和线性变换T上使用

2.4 特征值与特征向量

特征值与特征向量

$$A\xi=\lambda_0\xi$$

- 恒等变换：特征值都是1，零变换：特征值都为零
- 特征向量在线性变换作用下保持共线, 即在同一直线上
- 若$T(\xi)=\lambda_0\xi$，则属于特征值λ_0的特征向量ξ的线性组合(非零的$k_1\xi_1+\cdots+k_n\xi_n$)，也是T属于特征值λ_0的特征向量
- 若$\xi\in N(T)$且向量非零，则ξ是属于0的特征向量

求解特征值

- 矩阵求解特征值时构造$|\lambda I-A|$的特征多项式
- **定理1**：
  - 矩阵的迹$\mathrm{tr}(A)=\sum\lambda_i=\sum a_{ii}$是所有特征值之和
  - 矩阵的行列式$|A|=\prod\lambda_i$是所有特征值之积
- **定理3**：矩阵A属于不同特征值的特征向量线性无关
- 对于一个r阶的**幂等矩阵**，有r个特征值为1，$n-r$个特征值为0

- 幂零矩阵$A^k=0$的特征值只能是0

特征子空间

对于n阶复方阵$A \in \mathbb{C}^{n\times n}$，每个特征值λ对应的特征子空间：

$$E(\lambda)=\{\mathbf{x}\in\mathbb{C}^n|(A-\lambda\mathbf{I})\mathbf{x}=0\}$$

几何重数与代数重数

- 特征子空间的维数$\dim E(\lambda)$，即为特征值λ的**几何重数**

$$\dim E(\lambda)=n-\mathrm{rank}(A-\lambda\mathbf{I})$$

- λ的**代数重数**，即为特征方程根的重数
- **定理2**：**对于矩阵的同一特征值：几何重数$\leq$代数重数**（代数重要）

2.5 酉变换与酉矩阵

正交(酉)变换与正交(酉)矩阵

- 欧式空间：实内积空间
- 酉空间：复内积空间

欧式（酉）空间的线性变换保持向量内积不变，即称之为正交（酉）变换

$$(T(x),T(y))=(x,y),\forall x,y\in V$$

正交矩阵：n阶实方阵满足$A^TA=I$或$AA^T=I$

酉矩阵：n阶实方阵满足$A^HA=I$或$AA^H=I$

正交(酉)变换与正交(酉)矩阵的性质

定理1：正交（酉）变换等价判定定理，$T\in L(V)$以下四个命题等价

- T是正交（酉）变换
- T的长度保持不变$\|T(x)\|=\|x\|$
- ξ_i为V下的标准正交基，那么$T(\xi_i)$也是V中的标准正交基
- T在V的任意标准正交基下的矩阵为正交（酉）矩阵

命题1：正交（酉）矩阵A的性质

- 正交矩阵行列式为±1，酉矩阵的行列式模值为1
- $A^{-1}=A^T$是正交矩阵；$A^{-1}=A^H$是酉矩阵
- 正交（酉）矩阵的乘积仍为正交（酉）矩阵
- A的所有特征值的模值为1

定理2：A是正交（酉）矩阵当且仅当$A=[\alpha_1,\alpha_2,\cdots,\alpha_n]$构成的$n$个列向量$\alpha_i$构成欧式（酉）空间上的一组标准正交基。

Givens矩阵

Givens矩阵/初等旋转矩阵定义为，其本身为一个n阶的单位矩阵，然后在i,j的两个行列位置嵌入一组旋转$\cos\theta$与$\sin\theta$:

$$T(i,j,\theta)=\begin{bmatrix}1 & & & & & & \\ & \ddots & & & & & \\ & & \cos\theta & & \sin\theta & & \\ & & & \ddots & & & \\ & & -\sin\theta & & \cos\theta & & \\ & & & & & \ddots & \\ & & & & & & 1\end{bmatrix}$$

$$T(2,3,\theta)=\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\theta & \sin\theta\\ 0 & -\sin\theta & \cos\theta\end{pmatrix}$$

Givens矩阵有以下性质：

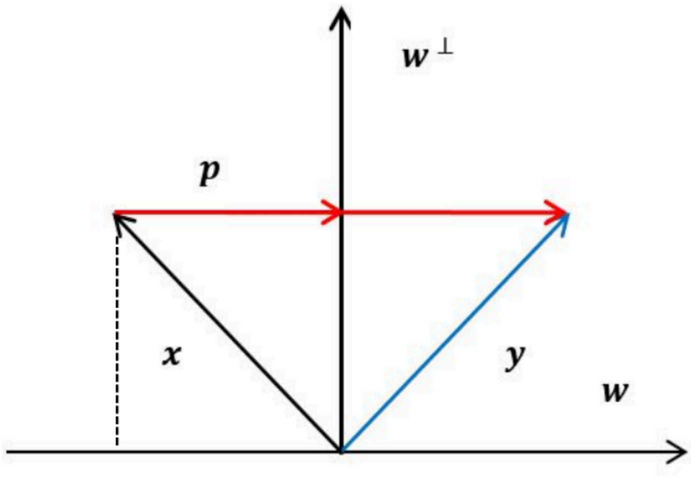
- Givens矩阵是正交矩阵 $T(i,j)^{-1}=T(i,j)^T$
- 若$y=T(i,j)x$，那么：
  - $y_k=x_k,k\neq i,j$
  - $y_i=\cos\theta x_i+\sin\theta x_j$
  - $y_j=-\sin\theta x_i+\cos\theta x_j$
  - 此性质即表明，若$x_i^2+x_j^2\neq 0$，则：$y_i=\sqrt{x_i^2+x_j^2},y_j=0$
- 可利用Givens矩阵在向量中介入零，使得$T(1,n)T(1,n-1)\cdots T(1,2)=\|x\|e_1$

Householder矩阵

Householder又称初等反射矩阵，有以下定义：

$$H(w)=I-2ww^H$$

本质的含义是将任意x沿着$x^{\perp}$方向进行对称反射，如图所示：



- $x+2p=y$
- $p=x-\text{Proj}_W^{\perp}x=-\text{Proj}_W x=-(x,w)w$
- $y=x-2(x,w)w=x-2w^Hxw=(I-2ww^H)x=H(w)x$

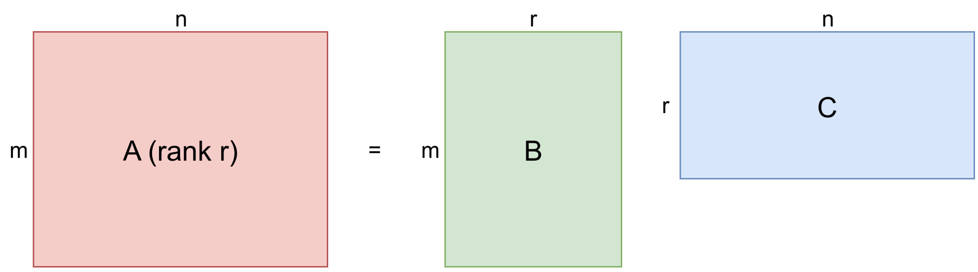
Householder矩阵有以下性质：

- Householder矩阵是一个酉矩阵：$H(w)^H=H(w)^{-1}=H(w)$
- Householder矩阵有$n-1$个特征值为1，有1个特征值为-1
- $y=H(w)x$若需要满足：$x^Hx=y^Hy, x^Hy=y^Hx$，可取:

$$w=\frac{e^{i\theta}}{\|x-y\|}(x-y)$$

第三章：矩阵分解

3.1 满秩分解



满秩分解

$$A=BC \quad A\in\mathbb{C}_r^{m\times n} \quad B\in\mathbb{C}_r^{m\times r}, C\in\mathbb{C}_r^{r\times n}$$

为了构建B，可以通过选取A中列向量构成的基底，或者其他方式挑选基，C即为选取基底下列向量的坐标

若两方阵乘积为单位阵，则这两矩阵必可逆

左逆和右逆

假设存在矩阵$A \in \mathbb{C}_r^{m\times n}$

- 左逆B存在$BA=I$，充分必要条件：A为列满秩矩阵
- 右逆B存在$AB=I$，充分必要条件：A为行满秩矩阵
- “左列右行“

补充：可逆矩阵的判断

- 可逆矩阵必为方阵
- 可逆矩阵行列式不为零$\det A \neq 0$
- 可逆矩阵必须为满秩矩阵
- 可逆矩阵需要满足$Ax=b$有唯一解，或者$Ax=0$只有零解
- 可逆矩阵的特征值都不为零
- 对角矩阵$\mathrm{diag}\{\lambda_1,\cdots\lambda_n\}$需要满足对角元素$\lambda_i \neq 0$，其逆矩阵为$\mathrm{diag}\{\lambda_1^{-1},\cdots\lambda_n^{-1}\}$
- 正交矩阵$A^{-1}=A^T$一定可逆

矩阵秩的重要结论

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

在上述条件下可以发现：$Ax=0$，$A^HAx=0$，$x^HA^HAx=0$有相同的解

补充：相抵分解

若矩阵$\mathbf{A}$与$\mathbf{B}$相抵：

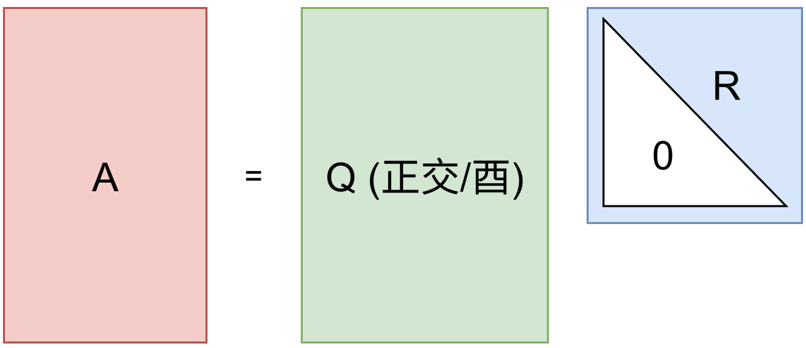
- 存在可逆矩阵$\mathbf{P}$与$\mathbf{Q}$，使得$\mathbf{A}=\mathbf{PBQ}$
- $\mathrm{rank}(\mathbf{A})=\mathrm{rank}(\mathbf{B})$
- 矩阵$\mathbf{A}$与$\mathbf{B}$可以通过有限次初等行变换得到同一个矩阵

矩阵$\mathbf{A}$的相抵分解：

$$\mathbf{A}=\mathbf{PBQ}=\mathbf{P}\begin{bmatrix}\mathbf{I}_r & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\mathbf{Q}$$

其中，$\mathbf{A}\in F^{m\times n}$, $\mathbf{P}$ 与 $\mathbf{Q}$ 分别为m与n维的方阵。其中$\begin{bmatrix}\mathbf{I}_r & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}$被称为矩阵$\mathbf{A}$的**相抵标准型**。

3.2 QR分解



QR分解

$$A = QR$$

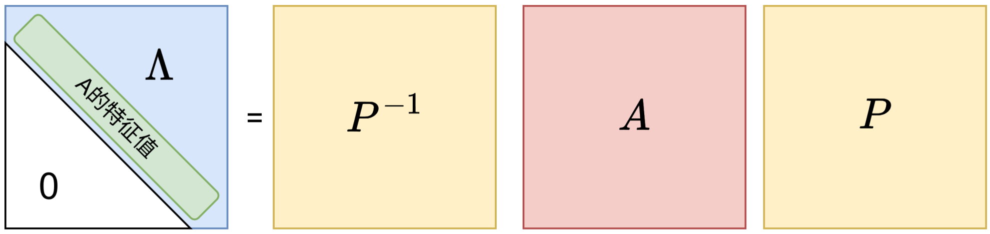
- Q 为酉矩阵，R为上三角矩阵
- 若在实数域中也存在这个满足的条件，那么称其为**正交三角分解**
- QR 分解可以直接由Gram-Schmidt正交化方法直接得到
- 实方阵
  - 若实方阵A满秩，则存在**唯一的QR分解**，R为**正线上三角阵**
  - 若不要求上三角阵R的**对角元素全为正实数,** 则导致矩阵A的**QR分解不唯一**
  - 若一实方阵既是正交矩阵又是正线上三角矩阵，则该矩阵一定是单位矩阵
- 复方阵
  - 若**复方阵**A**可逆**，则存在酉矩阵Q及正线上三角阵R 满足QR分解唯一

列/行正交规范矩阵

$$Q^H Q = I$$

- Q为列正交规范矩阵；Q^H为行正交规范矩阵

3.3 Schur分解



可逆矩阵P可为酉矩阵，若A的特征值都为实数，则为也可为正交矩阵

Schur分解

- 任意复方阵A相似于上三角阵Λ，Λ的对角元素为A的特征值：

$$\Lambda = P^{-1}AP$$

对上述的可逆矩阵P做QR分解得到：$P = UR$，即可推得下述的结论：

- 任意复方阵A酉相似于上三角阵Λ，U为酉矩阵:

$$\Lambda = U^H AU$$

- 对于一个实矩阵A，由于其特征值有可能是复数，因此，不一定存在正交矩阵Q，满足$\Lambda = Q^T AQ$
- 若矩阵A所有特征值为实数，那么可以满足：

$$\Lambda = Q^T AQ$$

Hamilton-Cayley定理

$$\text{矩阵的特征多项式为} f_A(\lambda) = |\lambda I - A|，\text{则} f_A(A) = 0$$

矩阵多项式

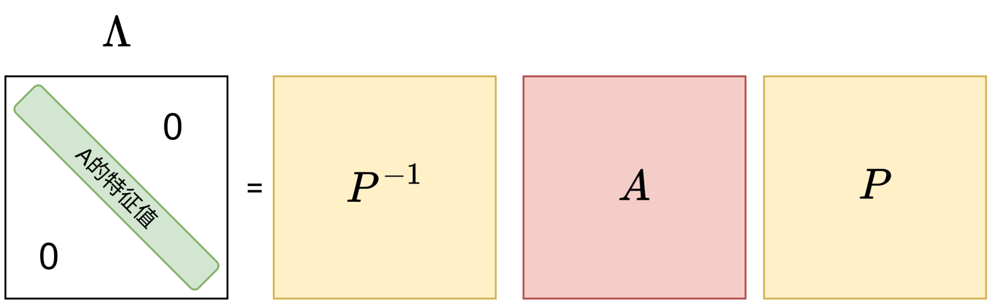
- **矩阵多项式**即将矩阵替换原先的λ

$$A \to \phi(\lambda) = \sum a_i \lambda^i \to \phi(A) = \sum a_i A^i$$

- **特征多项式**：即$f(A) = \det(\lambda I - A)$
- **零化多项式**：即$g(A) = 0$的多项式
- **最小多项式**：A的零化多项式中最小次数的首1多项式
- 最小多项式是唯一的，可整除矩阵A的任意零化多项式
- 最小多项式和经典凯莱-哈密顿定理仅适用于**方阵**的分析
- 对于n阶实对称矩阵的**代数重数 = 几何重数**，特征值λ的几何重数为：$n - \text{rank}(A - \lambda I)$
- 可以利用矩阵多项式求逆：$f_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots a_1\lambda + a_0$，那么：

$$A^{-1} = -\frac{1}{a_0}(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I)$$

3.4 对角化分解



单纯矩阵

- **单纯矩阵**A可通过可逆矩阵P变换到对角矩阵Λ，即**对角化分解**：

$$P^{-1}AP = \Lambda$$

则矩阵A为可**对角化矩阵=单纯矩阵**，实际上为Schur分解的特殊形式

- **单纯矩阵的判定方法**：
    - A有n个线性无关的特征向量
    - 每个特征值的**代数重数=几何重数**
    - $\sum_{i=1}^m \text{dim}E(\lambda_i) = n$
    - 最小多项式$m_A(\lambda)$无重根

其中$\text{dim}E(\lambda_i)$为特征子空间维数

- **几个单纯矩阵的例子**：
- $A^2 = A$幂等矩阵，即将其化为：$\lambda(\lambda - 1) = 0$

- 矩阵$A \in \mathbb{C}^{n\times n}$若所有特征值都互异

正规矩阵

- **正规矩阵**A可通过酉矩阵U变换到对角矩阵Λ，即**酉相似对角化分解**：

$$U^H AU = \Lambda$$

酉相似对角化是对角分解的一种特殊形式。

- **正规矩阵**都是可酉相似对角化的，这是**可酉相似对角化的**充要条件，满足以下条件：

$$A^H A = AA^H$$

- 常见的正规矩阵：
  - 对称矩阵$A = A^T$ **必须实数域**
  - 反对称矩阵$A = -A^T$ **必须实数域**
  - Hermite矩阵$A = A^H$
  - 反Hermite矩阵$A = -A^H$
  - 正交矩阵$A^T A = AA^T = I$ **必须实数域**
  - 酉矩阵$A^H A = AA^H = I$
- 复方阵A是正规矩阵当且仅当有n个特征向量构成$\mathbb{C}^n$ 空间的一组标准正交基，且属于A的不同特征值的特征向量正交。
- A, B矩阵**酉相似的充分必要条件是具有相同的特征值**
- 若矩阵$AB = BA$，则两个矩阵必有公共特征向量
- 对于任意$A \in \mathbb{C}^{m\times n}$:
  - $\operatorname{rank}(A) = \operatorname{rank}(A^H A) = \operatorname{rank}(AA^H)$
  - $A^H A$与AA^H的特征值均为非负实数
  - $A^H A$与AA^H的非零特征值相同

3.5 谱分解

正规矩阵的谱分解

- 一个正规矩阵$A \in C^{n\times n}$有$\lambda_1 \cdots \lambda_m$个不同的特征值
- 每个特征值λ_j的代数重数分别为$d_1 \cdots d_m$
- 每个特征值λ_j有d_j个单位正交的特征向量$\mathbf{u}_{j1} \cdots \mathbf{u}_{jd_j}$
- 矩阵A的谱分解式为：

$$A = \sum_{j=1}^{m} \lambda_j E_j = \sum_{j=1}^{m} \lambda_j \cdot \sum_{i=1}^{d_j} \mathbf{u}_{ji}\mathbf{u}_{ji}^H$$

$$A = U\Lambda U^H = \begin{bmatrix} \alpha_1 & \cdots & \alpha_n \end{bmatrix}_{n\times n} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} \alpha_1^H \\ \vdots \\ \alpha_n^H \end{bmatrix}$$

- 正规矩阵的谱阵有以下性质($i, j = 1, \cdots, m$且$i \neq j$):
  - $A = \sum \lambda_i E_i$
  - $E_i = E_i^H = (E_i)^2$
  - $E_i E_j = 0$
  - $E_i A = AE_i = \lambda_i E_i$
  - $\sum E_i = I$
  - 谱阵的集合$\{E_i\}$唯一

幂等矩阵与投影矩阵

- 幂等矩阵（投影矩阵）：$A = A^2$
- 正交投影矩阵：$A^2 = A = A^H$
- 幂等矩阵的性质：
  - A^H,A^*,$I - A$都是幂等矩阵
  - A为单纯矩阵且相似于对角阵$\Lambda = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$
  - $tr(A) = \operatorname{rank}(A)$
  - $N(A) = R(I - A)$
  - $\mathbb{C}^n = R(A) \dot{+} N(A)$
  - 若$Ax = x$则$x \in R(A)$
- 对于一个r阶的幂等矩阵，有r个特征值为1，$n - r$个特征值为0
- 对于任意n阶矩阵$\operatorname{rank} A^k, k \geq n$都相等稳定了

单纯矩阵谱分解

- 相比于正规矩阵，单纯矩阵的谱分解是更一般的形式
- 正规矩阵的谱阵的格式为：

$$E_j = \sum_{i=1}^{d_j} \mathbf{u}_{ji} \mathbf{u}_{ji}^H$$

- 而单纯矩阵的谱阵的格式为：

$$E_j = \sum_{i=1}^{d_j} \alpha_{ji} \beta_{ji}^H$$

- 单纯矩阵的谱阵是幂等矩阵，但不是正交投影矩阵
- 单纯矩阵的谱阵可直接通过特征多项式求解：

$$E_i = \frac{f_i(A)}{f_i(\lambda_i)} = \frac{\prod_{l \neq i} (A - \lambda_l I)}{\prod_{l \neq i} (\lambda_i - \lambda_l)}$$

3.6 Jordan分解

λ矩阵

- λ矩阵的定义
  - 以λ为元素的矩阵称为λ矩阵
  - 数字矩阵是特殊的λ矩阵
- λ矩阵的性质
  - λ矩阵的非零子式的最高阶数r为$A(\lambda)$的秩
  - $\lambda I-A$总是满秩的
  - $A(\lambda)$可逆的充分必要条件是$|A(\lambda)|$为非零常数
- λ矩阵相抵
  - 经过有限次初等变换（包括列变换）变化，使得$A(\lambda)=B(\lambda)$，则称$A(\lambda)$与$B(\lambda)$相抵，记为：

    $$A(\lambda)\cong B(\lambda)$$

    两个相抵的λ矩阵只能相差一个常数
  - 相抵的λ矩阵具有**相同的秩、各阶行列式因子、不变因子**
  - 若两个矩阵相抵，必须具有相同的Smith标准型
- **复方阵$A\sim B$，即两矩阵相似时，充分必要条件为的特征矩阵$(\lambda I-A)\cong(\lambda I-B)$ 相抵**

三个矩阵因子

- **行列式因子D_i**：矩阵第i阶所有的非零子式的最大公因子

  $$\begin{bmatrix}\lambda & \lambda \\ 0 & \lambda(\lambda-1)\end{bmatrix}$$

  上述矩阵有2阶，对于1阶子式，有：$|\lambda|,|\lambda|,|0|,|\lambda(\lambda-1)|$四个，对于非零子式的最大公因子为$D_1=\lambda$。对于2阶子式，有$\det\begin{bmatrix}\lambda & \lambda \\ 0 & \lambda(\lambda-1)\end{bmatrix}=\lambda^2(\lambda-1)$，因此$D_2=\lambda^2(\lambda-1)$
- **不变因子d_i**：

  $d_1=D_1,d_2=D_2/D_1,\cdots d_n=D_n/D_{n-1}$
- **初等因子组**：即剔除所有1次项目的不变因子，写在一块

  $d_1=1,d_2=\lambda,d_3=\lambda(\lambda+1)$

  那么这一组的初等因子为：$\lambda,\lambda,\lambda+1$

  注意，初等因子组存在顺序之分

Smith标准型

$A(\lambda)$的秩为r，那么其Smith标准形为：

$$A(\lambda) \cong \begin{bmatrix} d_1(\lambda) & & \\ & \ddots & \\ & & d_r(\lambda) \\ & & \end{bmatrix}$$

λ矩阵的Smith标准型是唯一的，其不一定是方阵

Jordan分解

- Jordan块

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i \end{bmatrix}_{n_i \times n_i}$$

  Jordan块的最小多项式为其特征多项式，即为：$(\lambda - \lambda_i)^{n_i}$

- Jordan分解

  给定复方阵A的特征矩阵$\lambda I - A$对应的初等因子为：

$$(\lambda - \lambda_1)^{n_1}, (\lambda - \lambda_2)^{n_2}, \cdots, (\lambda - \lambda_s)^{n_s}$$

  其对应的Jordan块为$J_1, \ldots J_s$那么A的Jordan标准型为：

$$J = \mathrm{diag}(J_1, \ldots, J_s)$$

例3.6.11 设A的特征矩阵

$$(\lambda I - A) \cong \mathrm{diag}(\lambda^2, \lambda(\lambda+1)^2, 1,1,1)$$

求A的Jordan标准形.

解: $(\lambda I - A)$初等因子为$\lambda^2, \lambda, (\lambda+1)^2$. 作Jordan块

$$J_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, J_2 = [0], J_3 = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}$$

$$J = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

  Jordan分级可以视为对对角分解的广义推广，在现代控制理论中有重要意义。

- **Frobenious定理**：复方阵A的特征矩阵$\lambda I - A$ 的Smith标准型为$\mathrm{diag}(d_1(\lambda) \ldots d_n(\lambda))$，$A$的最小多项式$m_A(\lambda) = d_n(\lambda)$

第四章：矩阵分析

4.1 向量范数

向量范数（3条）

向量范数的存在满足三条范数的公理：

- 正定性：$\parallel x \parallel \geq 0$等于零的情况仅在$x=0$
- 齐次性：$\|kx\| = |k| \parallel x \parallel$
- 三角不等式：$\|x+y\| \leq \parallel x \parallel + \|y\|$ 即三角形两边之和大于第三边

满足上述性质的$\parallel x \parallel$为向量x的范数，V是数域上的**赋范线性空间**

- 内积空间是赋范线性空间
- 在同一线性空间中可定义不同的范数

基本的向量范数

- 1范数 $\parallel x \parallel_1 = \sum |x_i|$
- 2范数 $\parallel x \parallel_2 = (\sum |x_i^2|)^{\frac{1}{2}}$
- ∞范数 $\parallel x \parallel_\infty = \max |x_i|$
- p范数 $\parallel x \parallel_p = (\sum |x_i|^p)^{\frac{1}{p}}$

范数的本质：将空间中的向量映射为一个非负实数的函数

不同空间的范数定义的单位圆是有区别的，但是都是有界闭集

线性空间V中任一范数$\parallel x \parallel$都是其坐标的连续函数

范数等价

范数等价：$k_1 \parallel x \parallel_\beta \leq \parallel x \parallel_\alpha \leq k_2 \parallel x \parallel_\beta \qquad k_1, k_2 > 0$

有限维线性空间中的任意向量范数都是等价的

范数等价的性质：

- 自反性：$1 \cdot \parallel x \parallel_\alpha \leq \parallel x \parallel_\alpha \leq 1 \cdot \parallel x \parallel_\alpha$
- 对称性：$1/k_1 \cdot \parallel x \parallel_\alpha \leq \parallel x \parallel_\beta \leq 1/k_2 \cdot \parallel x \parallel_\alpha$
- 传递性：β与γ范数等价，那么α同样和γ等价

4.2 矩阵范数

矩阵的向量范数

满足的条件与向量范数一样，不过将内容从向量换成了矩阵：

- 正定性：$\|A\| \geq 0$等于零的情况仅在$A=0$
- 齐次性：$\|kA\| = |k| \, \|A\|$
- 三角不等式：$\|A+B\| \leq \|A\| + \|B\|$

基本的矩阵向量范数

- $\|A\|_{v1}=\sum\sum|a_{ij}|$
- $\|A\|_{v2}=(\sum\sum|a_{ij}^2|)^{\frac{1}{2}}$
- $\|A\|_{v\infty}=\max|a_{ij}|$
- $\|A\|_{vp}=(\sum\sum|a_{ij}|^p)^{\frac{1}{p}}$

任意矩阵的向量范数$\|A\|$都是其自身元素的连续函数

范数等价：$k_1\|A\|_{v\beta}\leq\|A\|_{v\alpha}\leq k_2\|A\|_{v\beta}\qquad k_1,k_2>0$

线性空间$\mathbb{C}^{m\times n}$中的任意向量范数都是等价的

矩阵范数（3+1）

矩阵范数出了满足先前的三条性质外，还需要额外加上一条**相容性**，兼顾矩阵的乘法运算：

- 正定性：$\|A\|\geq 0$等于零的情况仅在$A=0$
- 齐次性：$\|kA\|=|k|\,\|A\|$
- 三角不等式：$\|A+B\|\leq\|A\|+\|B\|$
- 相容性：$\|AB\|\leq\|A\|\,\|B\|$

矩阵乘法相容性的不等式实际上保证了矩阵幂级数的收敛性，$\|A\|\leq 1$那么当$k\to\infty$，$\left\|A^k\right\|\to 0$

矩阵F范数

对于$\|A\|_{v2}$，为矩阵的Frobenious范数

$$\|A\|_{v2}=(\sum\sum|a_{ij}^2|)^{\frac{1}{2}}=\sqrt{\operatorname{tr}(A^HA)}=\|A\|_F$$

设U,V为两个酉矩阵，那么F范数满足：

- $\|UA\|_F=\|AV\|_F=\|UAV\|_F=\|A\|_F$，即为F范数的不变性
- 若A矩阵按列分块，记为$A=[\beta_1,\cdots,\beta_n]$那么：$\|A\|_F^2=\sum\|\beta_i\|_2^2$，矩阵$A$按行分块也有相同的结果
- $\|Ax\|\leq\|A\|_F\shortparallel x\shortparallel_2$

4.3 相容范数

向量范数与矩阵范数相容

向量范数$\shortparallel x\shortparallel_v$与矩阵范数$\|A\|_m$相容，即满足：

$$\|Ax\|_v\leq\|A\|_m\shortparallel x\shortparallel_v$$

- 向量2范数与矩阵F范数相容

定理1：若$\|A\|_m$是$\mathbb{C}^{n\times n}$的矩阵范数，则必存在$\mathbb{C}^n$上相容的向量范数

- 可以通过构造$\shortparallel x\shortparallel_v=\left\|x\alpha^T\right\|_m$并验证矩阵范数的四条性质证明

算子范数（诱导范数）

定理2： $\lVert x \rVert_v$为$\mathbb{C}^n$上的向量范数，对于$A \in \mathbb{C}^{m\times n}$：

$$\|A\| = \max_{\lVert x \rVert_v=1} \|Ax\|_v$$

那么$\|A\|$是一个和$\lVert x \rVert_v$相容的矩阵范数，称之为从属于$\lVert \cdot \rVert_v$的算子范数，或诱导范数

常见的算子范数

定理3： 从属于向量1,2,无穷范数的算子范数为：

- $\|A\|_1 = \max \sum_i |a_{ij}|$ 列和范数
- $\|A\|_\infty = \max \sum_j |a_{ij}|$ 行和范数
- $\|A\|_2 = \sqrt{\lambda_{\max}(A^H A)} = \sigma_{\max} A$ 谱范数

$\|A\|$不一定等于$\|A^T\|$

- $\|A\|_1 = \|A^T\|_\infty$
- $\|A\|_2 = \|A^T\|_2$

相容于同一个**向量范数**的矩阵范数有多个，例如F范数和矩阵2范数相容于向量2范数

常用不等式

Cauchy-Schwarz 不等式：**内积空间中两个向量内积的绝对值，不超过两个向量各自范数的乘积。**

- $|(\mathbf{x}, \mathbf{y})| \leq \lVert \mathbf{x} \rVert \, \|\mathbf{y}\|$
- $(\sum uv)^2 \leq (\sum u^2)(\sum v^2)$
- $\|Ax\|_v \leq \|A\|_m \lVert x \rVert_v$

定理4：$\|A\|$是某矩阵范数，$\|A\| < 1$，且$I - A$非奇异：

$$\left\|(I-A)^{-1}\right\| \leq \frac{\|I\|}{1 - \|A\|}$$

注：$\|I\|$不一定为1,可能大于1,但是算子范数等于1

4.4 特征值估计

谱和谱半径

复方阵特征值构成的集合称为谱，其模的最大值称之为谱半径

定理1：复方阵的谱半径不大于它的任一矩阵范数

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

定理2： 必存在某个矩阵范数使得：$\|A\| \leq \rho(A) + \epsilon$

盖尔圆盘

盖尔圆盘的中心在矩阵的对角元素上，其半径为每行非中心元素的模的和

$$A=\begin{bmatrix}1 & 0.02 & 0.11\\ 0.01 & i & 0.14\\ 0.02 & 0.01 & 0.5\end{bmatrix}$$

上述矩阵的三个圆盘的中心分别为$1, i, 0.5$，半径分别为$0.13, 0.15, 0.03$

定理3：矩阵A的任意一个特征值在其所有盖尔圆的并集区域内。

$$\lambda \in \cup_i G$$

定理4：若一个盖尔圆是孤立的，则这个独立的圆内仅有一个特征值；k个盖尔圆形成一个联通的区域，那么这个联通的区域内有k个特征值

- 由于A^T与A的特征值相同，且具有相同的圆形，因此可以做两组盖尔圆从而进一步缩小特征值的可能位置
- 并非每个盖尔圆内恰好有一个特征值

若盖尔圆的覆盖区域不含有原点，那么这个矩阵必然不含有特征值0，那么他一定是**非奇异矩阵**。

- 在这个情况下，若**矩阵是对角占优矩阵**，无论是列占优或行占优，即盖尔圆的圆心始终远离原点的距离大于其半径，那么矩阵一定非奇异。

若矩阵的所有盖尔圆两两互不相交, 则其有n个互不相同的特征值，其必然是是单纯矩阵

在使用盖尔圆估计矩阵的特征值时, 我们总希望获得更多的孤立圆，可采取对角变换的方法：

$$B=DAD^{-1}, \quad D=\operatorname{diag}(d_1,\cdots,d_n)$$

- 选取$d_i<1$，其余的元素为1，第i个盖尔圆会缩小，其余的都会放大
- 选取$d_i>1$，其余的元素为1，第i个盖尔圆会放大，其余的都会缩小

实矩阵的复特征值成对共轭出现

4.5 矩阵级数

向量序列按范数收敛

对于某个向量序列$\{x_i\}$，若存在向量x满足：

$$\lim_{k\to\infty}\|x_k-x\|_a=0$$

则称向量序列$\{x_i\}$按范数$||\cdot||_a$收敛于x，记作：

$$\lim_{k\to\infty}x_k=x$$

定理1：有限维空间中范数收敛是等价的

向量序列按坐标收敛

对于某个向量序列$\{x_i\}$，在其定义的赋范线性空间中V的一组基ϵ_i下有坐标：

$$\xi_k = [\xi_1^{(k)} \cdots \xi_n^{(k)}]^T$$

若满足$\lim_{k\to\infty} \xi_i^{(k)} = \xi_i$，则称向量序列按坐标收敛于向量$x$

注，这里的上标(k)代表向量序列的第k个$x_i, i = k$，不是平方，假设有k个x，这组基有n个，那么这样有一个平面的$k \times n$个数，按照坐标收敛需要满足随着序列的前进，每个坐标收敛到某个常数，产生n个稳定的值。

定理2：向量序列按范数收敛当且仅当它按坐标收敛

矩阵序列按坐标收敛

即矩阵序列中的每一个元素a_{ij}都有一个极限$a_{ij}^{(0)}$，则称矩阵序列按元素/坐标收敛：

$$\lim_{k\to\infty} A_k = A_0$$

若两个矩阵序列$\{A_k\}\{B_k\}$按坐标收敛于A, B，那么有：

- $\lim_{k\to\infty}(c_1A_k + c_2B_k) = c_1A + c_2B$
- $\lim_{k\to\infty}(A_kB_k) = AB$
- 若A_k和A可逆$\lim_{k\to\infty} A_k^{-1} = A^{-1}$

矩阵序列$\{A_k\}$收敛于A_0的充分必要条件为：

$$\lim_{k\to\infty} \|A_k - A_0\| = 0$$

矩阵幂的收敛问题

若A的某一范数满足$\|A\| < 1$，那么有：

$$\lim_{k\to\infty} A^k = 0$$

但是这一结论反过来不成立。

定理3：$\lim_{k\to\infty} A^k = 0$的充分必要条件是$\rho(A) < 1$

矩阵级数

给定矩阵序列$\{A_k\}$，$\sum_{k=1}^{\infty} A_k$则为**矩阵级数**，$\sum_{k=1}^{N} A_k$称之为**矩阵级数的部分和**

若矩阵序列$\lim_{n\to\infty} S_N = S$，那么矩阵级数收敛且有和$S$

- 若矩阵级数收敛，矩阵$A \in \mathbb{C}^{m\times n}$中每个元素的级数，共$mn$个级数收敛
- 若矩阵级数收敛，$\lim_{k\to\infty} A_k = 0$
- 若矩阵级数发散，矩阵$A \in \mathbb{C}^{m\times n}$中每个元素的级数，共$mn$个级数，至少有一个级数发散

矩阵级数**绝对收敛**，即为其每个mn个级数都收敛

- 若矩阵级数绝对收敛，那么矩阵级数一定收敛；但是不代表矩阵级数收敛时，矩阵级数绝对收敛
- 对任意常矩阵P, Q,若矩阵级数$\sum A_k$(绝对)收敛，$\sum PA_kQ$也一定（绝对）收敛
- **定理4**：矩阵级数$\sum A_k$绝对收敛，仅当对任意矩阵范数，$\sum \|A_k\|$收敛

矩阵幂级数

$A^0 = I$，定义矩阵幂级数：

$$\sum_{m=0}^{\infty} c_m A^m$$

矩阵幂级数绝对收敛要求$\sum |c_m||A^m|$收敛

收敛半径：若$|z| < r$时$\sum c_m z^m$收敛，而$|z| > r$时不收敛，则r称为收敛半径

- 收敛半径公式，对于级数$\sum_{k=0}^{\infty} c_k x^k$：

$$R = \lim_{k \to \infty} \left| \frac{c_k}{c_{k+1}} \right|$$

Abel型定理

设复变量幂级数$\sum_{m=0}^{\infty} c_m z^m$的收敛半径为$r$，矩阵$A \in \mathbb{C}^{n \times n}$的谱半径为$\rho(A)$，则：

- $\rho(A) < r$时，$\sum_{m=0}^{\infty} c_m A^m$绝对收敛
- $\rho(A) > r$时，$\sum_{m=0}^{\infty} c_m A^m$发散

Neumann级数

$\sum_{m=0}^{\infty} A^m$收敛当且仅当$\rho(A) < 1$，此时：

$$\sum_{m=0}^{\infty} A^m = (I - A)^{-1}$$

注：可逆矩阵$(I - A)^{-1}$可与矩阵A的乘积交换：

$$(I - A)^{-1} A = A(I - A)^{-1}$$

4.6 矩阵函数

矩阵函数

给定幂级数$\sum_{m=0}^{\infty} c_m z^m$，假设其有一个收敛半径$r$，并收敛于$f(z)$

若复方阵A满足$\rho(A) < r$，称收敛的矩阵幂级数为**矩阵函数**：

$$f(A) = \sum_{m=0}^{\infty} c_m A^m$$

常见的矩阵函数

都是从$m = 0$求和到正无穷∞：

对于A都是复方阵$\forall A \in \mathbb{C}^{n \times n}$

- $e^A = \sum \frac{1}{m!} A^m$
- $\sin A = \sum \frac{(-1)^m}{(2m+1)!} A^{2m+1}$
- $\sin A = \sum \frac{(-1)^m}{(2m)!} A^{2m}$

同时需要$\rho(A) < 1$的情况：

- $(I-A)^{-1} = \sum A^m$
- $\ln(I+A) = \sum \frac{(-1)^m}{m+1} A^{m+1}$

矩阵函数的性质

复方阵情况下以下结论成立：

- $\cos(-A) = \cos(A), \sin(-A) = -\sin(A)$
- $e^{iA} = \cos A + i \sin A$
- $\cos A = \frac{1}{2}(e^{iA} + e^{-iA})$
- $\sin A = \frac{1}{2}(e^{iA} - e^{-iA})$

定理1： $AB = BA$，则$e^A e^B = e^B e^A = e^{A+B}$

- 若A是方阵：$e^A e^{-A} = e^{-A} e^A = I$，即$(e^A)^{-1} = e^{-A}$
- 若A是方阵：$\sin^2 A + \cos^2 A = I$

含参矩阵函数：

$$f(At) = \sum_{m=0}^{\infty} c_m (At)^m = \sum_{m=0}^{\infty} c_m t^m A^m, |t|\rho(A) < r$$

Sylvester公式

定理2：若$P^{-1}AP = B$，$f(A) = P^{-1} f(B) P$

对定理2的证明可以分为单纯矩阵和非单纯矩阵的形式，对于单纯矩阵的证明方法对角化矩阵A后较为容易得到，但是对于非单纯矩阵的证明，需要将A转化为标准Jordan形：$P^{-1}AP = J$，此时$A^k = PJ^kP^{-1}$，对于$J^k = \mathrm{diag}(J_1^k, \cdots, J_s^k)$，对于每个Jordan块，有：

$$J_i^k = \begin{bmatrix} \lambda_i^k & C_k^1 \lambda_i^{k-1} & C_k^2 \lambda_i^{k-2} & \dots & C_k^{n_i-1} \lambda_i^{k-n_i+1} \\ & \lambda_i^k & C_k^1 \lambda_i^{k-1} & \dots & C_k^{n_i-2} \lambda_i^{k-n_i+2} \\ & & \ddots & \ddots & \vdots \\ & & & \lambda_i^k & C_k^1 \lambda_i^{k-1} \\ & & & & \lambda_i^k \end{bmatrix}$$

$C_m^n = \frac{m!}{n! \times (m-n)!} = C_{m-1}^n + C_{m-1}^{n-1}$

上式即为Sylvester公式

Sylvester公式计算繁琐，解决方法是将矩阵分解为多项式的形式

定理3：若矩阵A的谱半径小于收敛半径r，那么必存在$l-1$阶矩阵多项式，其中l是矩阵A的最小多项式次数，使得：

$$f(A) = p(A) = \beta_0 I + \beta_1 A + \cdots + \beta_{l-1} A^{l-1}$$

谱上一致方法

谱上给定：某个复方阵A有s个互异的特征值λ_i，其最小多项式的形式为：

$$m_A(\lambda)=\prod(\lambda-\lambda_i)^{n_i}$$

若复函数$f(z)$和各阶导数$f^{(j)}(z)$在λ_i处有定义且连续，则称：

- $f(z)$在A的谱上给定
- λ_i为谱点，$f^{(i)}(\lambda_i)$为谱值

例4.6.6 考查函数$f(z)$在矩阵A的谱上是否有定义，其中

$$f(z)=\frac{1}{(z-3)(z-4)},\ A=\begin{bmatrix}2&0&0\\0&1&1\\0&0&1\end{bmatrix}$$

解：A的最小多项式为$m_A(\lambda)=(\lambda-2)(\lambda-1)^2$.

其谱点分别为$\lambda_1=2,\lambda_2=1$. 相应的谱值为

$$f(2)=\frac{1}{2},f(1)=\frac{1}{6},f'(1)=\frac{5}{36}.$$

所以$f(z)$在A的谱上有定义（在A的谱上给定）.

谱上一致：若函数$f(\lambda)$和$p(\lambda)$谱上给定且满足：

$$f^{(n_i)}(\lambda_i)=p^{(n_i)}(\lambda_i)$$

即他们的各阶导数相同，则称函数$f(\lambda)$和$p(\lambda)$谱上一致

定理4：$f(z)=\sum c_m z^m$和$p(z)=\sum\beta_i z^i$满足$f(A)=p(A)$的充分必要条件是$f(z)$和$p(z)$谱上一致