第1章：线性空间与线性变换

内容概要

- §1.1 线性空间 ～ 向量空间 R^n

  要点：空间的代数与几何结构，与向量空间R^n的关系

- §1.2 内积空间 ～ 向量空间 R^n

  要点：线性空间中向量的度量，R^n中的内积的推广

- §1.3 线性变换 ～ 矩阵空间 $R^{m\times n}$

  要点：线性空间之间的映射（线性映射），与矩阵空间$R^{m\times n}$的关系

1

第1章：线性空间与线性变换

预备知识

- 集合：笼统地说是指一些事物（或者对象，称为元素）组成的整体。

集合的表示：枚举、表达式，如

$$A=\{1,2,5\};\quad A=\{1,2,3,\ldots\};\quad A=\{x+x^2 \mid x>0\}$$

集合的运算：并（∪），交（∩）

2

第1章：线性空间与线性变换

预备知识

- 数域：一种数集，对四则运算封闭（除数不为零）。

比如有理数域、实数域（R）和复数域（C）。

实数域和复数域是工程上较常用的两个数域。

3

第1章：线性空间与线性变换

- 内容概述:
  - 线性空间的一般概念

    重点：空间的代数与几何结构，度量，与向量空间R^n的关系
  - 线性变换

    重点：矩阵处理方法，与矩阵的关系
- 特点:
  - 研究代数结构——具有线性运算的集合
  - 研究几何结构——空间的维数和基
  - 看重的不是研究对象本身，而是对象之间的结构关系
  - 研究的关注点：对象之间数量关系的矩阵处理
  - 学习特点：具有抽象性和一般性

4

1.1 线性空间

一、线性空间的概念

- n 维向量空间R^n (R^2，R^3): 结构, 表示, 运算, 度量等
- R^n到线性空间的推广思想：
  - 抽象出线性运算的本质，在任意研究对象的集合上定义具有线性运算的代数结构。

5

1.1 线性空间

一、线性空间的概念

- 线性空间的定义 (定义1.1 (P 1))

设V是一个非空集合，F是一个数域。

在V和F上定义了两种运算，加法与数乘:

$$\forall\, \alpha, \beta \in V, \alpha+\beta \in V;\ \forall\, \alpha \in V, k \in F, k\alpha \in V$$

并且满足以下8条运算性质:

- **5条运算律：加法交换律、结合律；乘法结合律；分配律1、2；**
- **3个特殊元素：零元, 负元, 数1**

则称V为数域F上的线性空间。
V中的元素称为向量。
F为实（复）数域时，称V为实（复）线性空间。

6

1.1 线性空间

一、线性空间的概念

- n 维向量空间R^n (R^2，R^3): 结构, 表示, 运算, 度量等
- R^n到线性空间的推广思想：
  - 抽象出线性运算的本质，在任意研究对象的集合上定义具有线性运算的代数结构。
- 线性空间的定义 (定义1.1 (P 1)) 数F域上的线性空间V
  - 三要素： V(F), 加法与数乘, 运算性质
    - 非空集合V与数域F
    - 两种运算：向量的加法和数乘向量 (可以不同于熟知的加法和乘法运算！)
    - 8条运算性质(公理定义), 含特殊元素(零元, 负元, 数1)

7

常见的线性空间

F=R或C

- $F^n = \{ X=(x_1, x_2, \ldots, x_n)^T: x_i \in F \}$

  运算：向量加法和数乘向量
  - R^n；C^n
- $F^{m\times n} = \{ A=[a_{ij}]_{m\times n}: a_{ij} \in F \}$;

  运算：矩阵的加法和数乘矩阵
  - $R^{m\times n}$；$C^{m\times n}$
- $P_n[x] = \{ p(x)=\sum_{i=0}^{n-1} a_i x^i : a_i \in R \}$ （<=n-1 阶多项式空间）

  运算：多项式的加法和数乘
- $C[a, b] = \{ f(x): f(x)$ 为 $[a, b]$ 上连续(实)函数$\}$

  运算：函数的加法和数乘 (还有$C^1[a, b]$等)
- 不一样的线性空间：$V=R^+, F=R, a\oplus b=ab, \lambda\otimes a=a^\lambda$

8

几点说明

- 线性空间不能离开某一数域来定义。

实际上，对于不同数域，同一个集合构成的线性空间会不同。

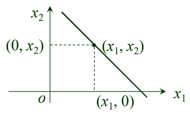
甚至一种能成为线性空间而另一种不能成为线性空间。

- 数域 中的运算是具体的四则运算，而 中所定义的加法运算和数乘运算则可以十分抽象。
- 封闭性的条件也很重要。

9

线性空间的判断

- 例子：次数为n-1的实多项式集合：
  $\{\, p(x)=\sum_{i=0}^{n-1} a_i x^i：\ a_i \in R，且\ a_{n-1}$不为零$\}$
- 平面上不过原点的直线上的点的集合，如V=$\{(x_1, x_2)：\ x_1, x_2 \in R，且\ x_1+x_2=1\}$；



- 空间中不过原点的直线(或平面)上的点的集合。

10

不是线性空间的例子与判别

- 例子：次数为n-1的实多项式集合：
  $\{\, p(x)=\sum_{i=0}^{n-1} a_i x^i：\ a_i \in R，且\ a_{n-1}$不为零$\}$；
  平面上不过原点的直线上的点的集合，如
  V=$\{(x_1, x_2)：\ x_1, x_2 \in R，且\ x_1+x_2=1\}$；
  空间中不过原点的直线(或平面)上的点的集合。
- 判别：依定义，即若不满足定义中的条件，则不是线性空间。如
  向量加法或数乘不封闭；
  不含特殊元素；……

11

线性空间的抽象

- 线性空间的一般形式：
  - V(F)，V中元素被统称为向量：$\alpha,\ \beta,\ \gamma,\ \ldots$
- 线性空间的简单性质（共性）：

定理1.1 V(F) 具有性质：

（1） V(F)中的零元素是惟一的。

（2） V(F) 中任何元素的负元素是惟一的。

（3）数零和零元素的性质：

数0 $0\alpha=0，k0=0；\ k\alpha=0 \Rightarrow \alpha=0$ 或 $k=0$

（4） $-\alpha=(-1)\alpha$ 向量0

$0\alpha=(0+0)\alpha=0\alpha+0\alpha \ \Rightarrow\ 0\alpha=0$

12

线性空间的抽象

- 线性空间的一般形式：
  - V(F)，V中元素被统称为向量：$\alpha,\ \beta,\ \gamma,\ \ldots$
- 线性空间的简单性质（共性）：

定理1.1 V(F) 具有性质：

（1） V(F)中的零元素是惟一的。

（2） V(F) 中任何元素的负元素是惟一的。

（3）数零和零元素的性质：

数0 $0\alpha=0，k0=0；\ k\alpha=0 \Rightarrow \alpha=0$ 或 $k=0$

（4） $-\alpha=(-1)\alpha$ 向量0

13

线性空间的抽象

- 线性空间的一般形式：
  - V(F)，V中元素被统称为向量：$\alpha,\ \beta,\ \gamma,\ \ldots$
- 线性空间的简单性质（共性）：

定理1.1 V(F) 具有性质：

（1） V(F)中的零元素是惟一的。

（2） V(F) 中任何元素的负元素是惟一的。

（3）数零和零元素的性质：

数0 $0\alpha=0，k0=0；\ k\alpha=0 \Rightarrow \alpha=0$ 或 $k=0$

（4） $-\alpha=(-1)\alpha$ 向量0

若 $k\alpha=0$, k不为0, 则 $\alpha=k^{-1}k\alpha=k^{-1}0=0$

14

线性空间的抽象

- 线性空间的一般形式：
  - V(F)，V中元素被统称为向量：$\alpha,\ \beta,\ \gamma,\ \ldots$
- 线性空间的简单性质（共性）：

定理1.1 V(F) 具有性质：

（1） V(F)中的零元素是惟一的。

（2） V(F) 中任何元素的负元素是惟一的。

（3）数零和零元素的性质：

数0 $0\alpha=0，k0=0；\ k\alpha=0 \Rightarrow \alpha=0$ 或 $k=0$

（4） $-\alpha=(-1)\alpha$ 向量0

$\alpha+(-1)\alpha=(1-1)\alpha=0\alpha=0 \ \Rightarrow\ -\alpha=(-1)\alpha$

15

二、线性空间的基和维数

- 向量的线性相关与线性无关：
  - 定义形式和向量空间R^n中的定义一样。
  - 有关性质与定理和R^n中的结果一样。

例题1 证明C[0, 1]空间中的向量组
$\{e^x，e^{2x}，e^{3x}\ldots，e^{nx}\}$，$x\in[0，1]$
线性无关。

令 $y=e^x\ (>0)$，考察多项式 $p(y)=\sum_{i=1}^{n} a_i y^i$

$y, y^2, \cdots, y^n$ 线性无关。

16

二、线性空间的基和维数

- 基与维数的概念(P 3)

定义1.2 设V(F)为线性空间，若存在一组线性无关的向量$\alpha_1，\alpha_2，\ldots，\alpha_n$，使得V中任一向量均可由其线性表示，即任给$\beta$，存在F中数组$\{x_1，x_2，\ldots，x_n\}$，使得 $\beta=x_1\alpha_1+x_2\alpha_2+x_3\alpha_3+\ldots+x_n\alpha_n$，
则称向量组$\{\alpha_1，\alpha_2，\ldots，\alpha_n\}$为V的一组基，基中所含向量的个数称为V的维数，记为 $\dim V=n,\ n<+\infty$，或 $n=+\infty$.

- 有用的记号：$\beta=\sum_{i=1}^{n} x_i\alpha_i=(\alpha_1,\alpha_2,\ldots,\alpha_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

17

二、线性空间的基和维数

- 常见线性空间的基与维数：
  - $F^n=\{X=(x_1，x_2，\ldots，x_n)^T：x_i\in F\}$，$\dim F^n=n$，
    $(x_1，x_2，\ldots，x_n)^T=\sum_{i=1}^{n} x_i e_i$，$e_i=(0, 0, \ldots, 1, 0, \ldots, 0)^T$，
    $\{e_1, e_2, \ldots, e_n\}$ 线性无关，构成一组基，称为自然基
  - $R^{m\times n}=\{A=[a_{ij}]: a_{ij}\in R\}$, 自然基$\{E_{ij}\}$, $\dim R^{m\times n}=m\times n$，
    $A=[a_{ij}]=\sum a_{ij}E_{ij}$，矩阵 E_{ij} 的第i行j列元素为1，其余元素为0。
  - 如 $R^{2\times 2}$ $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}=a_{11}E_{11}+a_{12}E_{12}+a_{21}E_{21}+a_{22}E_{22}$

18

二、线性空间的基和维数

- 常见线性空间的基与维数：
  - $P_n[x]$，自然基$\{1, x, x^2, x^3\dots, x^{n-1}\}$，$dim\ P_n[x]=n$
  - $P[x]$：（实或复）多项式全体，
    自然基$\{1, x, x^2, \dots, x^{n-1}\dots\}$，$dim\ P[x]=+\infty$
  - $C[a, b]$，$\{1, x, x^2, \dots, x^{n-1}\dots\}\subseteq C[a, b]$，$dim\ C[a, b]=+\infty$
  - 复数集C：作为C上的线性空间；作为R上的线性空间
    基与维数： 1，$dim\ C=1$； $\{1, i\}$，$dim\ C=2$

约定：

- $V_n(F)$表示数域F上的 n 维线性空间。
- 本课程主要研究有限维线性空间。

- 定理1.2 $V_n(F)$中任n个线性无关的向量均构成它的基。

19

三、坐标（向量在给定基下的线性表示）

1 定义1.3（P3）

设$\{\alpha_1, \alpha_2, \dots, \alpha_n\}$是线性空间$V_n(F)$的一组基，

$\forall\beta\in V_n(F)$，有 $\beta=\sum_{i=1}^{n}x_i\alpha_i$ 则称$x_1, x_2, \dots, x_n$是β在基$\{\alpha_i\}$下的坐标，

称$X=(x_1, x_2, \dots, x_n)^T$为$\beta$在基$\{\alpha_i\}$下的坐标向量，简称坐标。

$$\beta=\sum_{k=1}^{n}x_k\alpha_k=(\alpha_1,\alpha_2,\dots,\alpha_n)X=(\alpha_1,\alpha_2,\dots,\alpha_n)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

要点：坐标与基有关，向量在给定基下的坐标惟一；
坐标的表达形式：F^n中向量。

20

“坐标”的相关讨论

- 一般来说，线性空间及其元素是抽象的对象，不同空间的元素完全可以具有千差万别的类别及性质。但坐标表示把这种差别留给了基，由坐标所组成的新向量仅由数域中的数表示出来。
- 更进一步，原本抽象的“加法”及“数乘”经过坐标表示就演化为向量加法及数对向量的数乘。

21

“坐标”的意义讨论

1° $x+y=(\xi_1x_1+\xi_2x_2+\cdots+\xi_nx_n)+(\eta_1x_1+\eta_2x_2+\cdots+\eta_nx_n)$

$=(\xi_1+\eta_1)x_1+(\xi_2+\eta_2)x_2+\cdots+(\xi_n+\eta_n)x_n$

正对应 $\begin{cases}x=(\xi_1,\xi_2,\cdots,\xi_n)\\y=(\eta_1,\eta_2,\cdots,\eta_n)\end{cases}$ $\rightarrow x+y=(\xi_1+\eta_1,\xi_2+\eta_2,,\xi_n+\eta_n)$

2° $kx=k(\xi_1x_1+\xi_2x_2+\cdots+\xi_nx_n)=(k\xi_1)x_1+(k\xi_2)x_2+\cdots+(k\xi_n)x_n$

$\rightarrow(k\xi_1,k\xi_2,\cdots,k\xi_n)$

正对应 $x=(\xi_1,\xi_2,\cdots,\xi_n)$ $\rightarrow kx=(k\xi_1,k\xi_2,\cdots,k\xi_n)$

22

“坐标”的相关讨论

- 同一元素在不同基下坐标不同，后面会讨论坐标之间的具体关系。

23

三、坐标（向量在给定基下的线性表示）

- 例1 常见线性空间的基与坐标：
  - $F^n=\{X=(x_1, x_2, \dots, x_n)^T: x_i\in F\}$，$dim\ F^n=n$，
    $X=(x_1, x_2, \dots, x_n)^T=\sum_{i=1}^{n}x_ie_i$，$e_i=(0, 0, \dots, 1, 0, \dots, 0)^T$，
    X在自然基下的坐标就是自己：$(x_1, x_2, \dots, x_n)^T$
  - $R^{m\times n}=\{A=[a_{ij}]: a_{ij}\in R\}$，自然基$\{E_{ij}\}$，$dim\ R^{m\times n}=m\times n$，
    $A=[a_{ij}]=\sum a_{ij}E_{ij}$，于是A在自然基下的坐标就是A的元素构成的$mn$维向量。
  - 如 $R^{2\times2}$

$$(a_{11},a_{12},a_{21},a_{22})^T$$

$$\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}=a_{11}E_{11}+a_{12}E_{12}+a_{21}E_{21}+a_{22}E_{22}$$

24

三、坐标（向量在给定基下的线性表示）

- 常见线性空间的基与坐标（续）：
  - $P_n[x]$中多项式 $p(x)=\sum_{i=0}^{n-1}a_ix^i$
    在自然基$\{1, x, x^2, x^3\dots, x^{n-1}\}$下的坐标：$(a_0,a_1,\cdots,a_{n-1})^T$
  - $P[x]$中多项式 $p(x)=\sum_{i=0}^{m}a_ix^i$
    在自然基$\{1, x, x^2, \dots, x^n\dots\}$下的坐标：$(a_0,a_1,\cdots,a_m,0,\cdots)^T$
  - 复数集C：作为C上的线性空间；作为R上的线性空间
    基与坐标： 1，$z=z\cdot1$； $\{1, i\}$，$z=a+bi$

25

$$\beta=(\alpha_1,\alpha_2,\dots,\alpha_n)X$$

- 例2 设空间$P_4[x]$的两组基为：
  $\{1, x, x^2, x^3\}$和
  $\{1, (x-1)^1, (x-1)^2, (x-1)^3\}$
  求$f(x)=2+3x+4x^2+x^3$在这两组基下的坐标。
- 推广($(x-1)$ 变 $(x-x_0)$)?

归纳：

- 线性空间$V_n(F)$在任意一组基下的坐标属于F^n。
- 每一个常用的线性空间都有一组“自然基”，在这组基下，向量的坐标容易求得。
- 求坐标方法的各异性。

26

$$\beta=(\alpha_1,\alpha_2,\dots,\alpha_n)X$$

2 线性空间$V_n(F)$与F^n的同构

- $V_n(F) \xleftrightarrow{\text{坐标关系}} F^n$
  β 基$\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ X
- 由此，给定基，可建立一个一一对应关系σ
  $\forall\beta\in V_n(F)$，$\exists X\in F^n$，$\sigma(\beta)=X$
  - $\sigma(\beta_1+\beta_2)=\sigma(\beta_1)+\sigma(\beta_2)$ $\sigma(\alpha)=o\Leftrightarrow\alpha=o$
  - $\sigma(k\beta)=k\sigma(\beta)$ $\sigma(-\alpha)=-\sigma(\alpha)$

$$\beta_1=(\alpha_1,\alpha_2,\dots,\alpha_n)X=\sum_{k=1}^{n}x_k\alpha_k,\ \beta_2=(\alpha_1,\alpha_2,\dots,\alpha_n)Y=\sum_{k=1}^{n}y_k\alpha_k$$

$$\Rightarrow\ \beta_1+\beta_2=\sum_{k=1}^{n}(x_k+y_k)\alpha_k=(\alpha_1,\alpha_2,\dots,\alpha_n)(X+Y)$$

27

$\beta=(\alpha_1,\alpha_2,\ldots,\alpha_n)X$

2 线性空间$V_n(F)$与F^n的同构

坐标关系

- $V_n(F) \longleftrightarrow F^n$

  β 基$\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ X
- 由此，给定基，可建立一个一一对应关系σ

  $\forall \beta \in V_n(F)$，$\exists X \in F^n$， $\sigma(\beta)=X$
  - $\sigma(\beta_1+\beta_2)=\sigma(\beta_1)+\sigma(\beta_2)$ $\quad \sigma(\alpha)=o \Leftrightarrow \alpha=o$
  - $\sigma(k\beta)=k\sigma(\beta)$ $\quad \sigma(-\alpha)=-\sigma(\alpha)$

$$\beta_1=(\alpha_1,\alpha_2,\ldots,\alpha_n)X=\sum_{k=1}^n x_k\alpha_k,\ \beta_2=(\alpha_1,\alpha_2,\ldots,\alpha_n)Y=\sum_{k=1}^n y_k\alpha_k$$

$$\Rightarrow\ k\beta_1=\sum_{k=1}^n kx_k\alpha_k=(\alpha_1,\alpha_2,\ldots,\alpha_n)(kX)$$

$\beta=(\alpha_1,\alpha_2,\ldots,\alpha_n)X$

2 线性空间$V_n(F)$与F^n的同构

坐标关系

- $V_n(F) \longleftrightarrow F^n$

  β 基$\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ X
- 由此，给定基，可建立一个一一对应关系σ

  $\forall \beta \in V_n(F)$，$\exists X \in F^n$， $\sigma(\beta)=X$
  - $\sigma(\beta_1+\beta_2)=\sigma(\beta_1)+\sigma(\beta_2)$ $\quad \sigma(\alpha)=o \Leftrightarrow \alpha=o$
  - $\sigma(k\beta)=k\sigma(\beta)$ $\quad \sigma(-\alpha)=-\sigma(\alpha)$
- 在关系σ下，线性空间$V_n(F)$ 和 F^n 同构。

$$\sigma(k_1\beta_1+k_2\beta_2)=k_1\sigma(\beta_1)+k_2\sigma(\beta_2)\ \Rightarrow\ \sigma(\sum_{i=1}^m k_i\beta_i)=\sum_{i=1}^m k_i\sigma(\beta_i)$$

$\sigma(\sum_{i=1}^m k_i\beta_i)=\sum_{i=1}^m k_i\sigma(\beta_i)$

$\sigma(\alpha)=o \Leftrightarrow \alpha=o$

同构的性质

- 定理1.3 $V_n(F)$中向量$\{\beta_1, \beta_2, \ldots, \beta_m\}$线性相(无)关 ⇔ 它们在给定基下的坐标

  $\{X_1, X_2, \ldots, X_m\}$在$F^n$中线性相(无)关。
- 同构保持线性关系不变。
- 应用：

  借助于空间F^n中已经有的结论和方法研究一般线性空间$V_n(F)$中的线性关系。

- **例10(P5)** 讨论$P_4[x]$中向量$f_1=1+2x+4x^3$，$f_2=x+x^2+4x^3$，$f_3=1+x-3x^2$，$f_4=-2x+x^3$的线性相关性。(线性无关，构成一组基！)

解：取自然基，转化为坐标(系数)空间R^4中的问题。

- 例题 设$R^{2\times2}$中向量组$\{A_i\}$

$$A_1=\begin{bmatrix}1&1\\1&2\end{bmatrix}\quad A_2=\begin{bmatrix}0&2\\1&3\end{bmatrix}\quad A_3=\begin{bmatrix}3&1\\0&1\end{bmatrix}\quad A_4=\begin{bmatrix}2&-4\\-3&-7\end{bmatrix}$$

1 讨论$\{A_i\}$的线性相关性；
2 求向量组的秩和极大线性无关组；
3 把其余的向量表示成极大线性无关组的线性组合。

四、基变换和坐标变换

- 讨论：
  - 不同的基之间的关系
  - 同一个向量在不同基下坐标之间的关系
- 基变换公式

  设线性空间中有两组基：$\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ $\{\beta_1,\beta_2,\cdots,\beta_n\}$

  则 $\beta_i=(\alpha_1\ \alpha_2 \ldots \alpha_n)C_i, i=1,2,\cdots,n$

令 $C_{n\times n}=(C_1\ C_2 \ldots C_n)$ 则简记

$$(\beta_1\ \beta_2 \ldots \beta_n)=(\alpha_1\ \alpha_2 \ldots \alpha_n)C_{n\times n}$$

四、基变换和坐标变换

- 讨论：
  - 不同的基之间的关系
  - 同一个向量在不同基下坐标之间的关系
- 基变换公式

  设线性空间中有两组基：$\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ $\{\beta_1,\beta_2,\cdots,\beta_n\}$

  则 $(\beta_1\ \beta_2 \ldots \beta_n)=(\alpha_1\ \alpha_2 \ldots \alpha_n)C_{n\times n}$

过渡矩阵

过渡矩阵C的性质：
- C 的第i列是 β_i 在基$\{\alpha_k\}$下的坐标
- C 为非奇异矩阵（Th1.3）

$\sigma(\sum_{i=1}^n y_i\beta_i)=\sum_{i=1}^n y_i\sigma(\beta_i)$

2 坐标变换公式

已知
- 空间中两组基：$\{\alpha_1,\alpha_2,\ldots,\alpha_n\},\{\beta_1,\beta_2,\cdots,\beta_n\}$

满足： $(\beta_1\ \beta_2 \ldots \beta_n)=(\alpha_1\ \alpha_2 \ldots \alpha_n)C_{n\times n}$
- 设 $\alpha=(\alpha_1\ \alpha_2 \ldots \alpha_n)X$； $\alpha=(\beta_1\ \beta_2 \ldots \beta_n)Y$
- 讨论X和Y的关系

定理 1.4 $X=CY$ 或 $Y=C^{-1}X$

设由第一组基确定的同构映射为σ，则

$$X=\sigma(\alpha)=\sigma(\sum_{i=1}^n y_i\beta_i)=\sum_{i=1}^n y_i\sigma(\beta_i)=\sum_{i=1}^n y_iC_i=CY$$

2 坐标变换公式

如何求过渡矩阵？
- 从定义出发（坐标表示）；
- 对F^n,可通过矩阵求逆：由

$$(\beta_1\ \beta_2 \cdots \beta_n)=(\alpha_1\ \alpha_2 \cdots \alpha_n)C_{n\times n}$$

有 $C_{n\times n}=(\alpha_1\ \alpha_2 \cdots \alpha_n)^{-1}(\beta_1\ \beta_2 \cdots \beta_n)$

或 $C_{n\times n}^{-1}=(\beta_1\ \beta_2 \cdots \beta_n)^{-1}(\alpha_1\ \alpha_2 \cdots \alpha_n)$

定理1.4的应用：知C和Y，求$X=CY$；或知C和X，求$Y=C^{-1}X$。

例题11（P6）

- 例题 已知空间$R^{2\times2}$中两组基（I）$\{E_{ij}\}$

（II）$\left\{\begin{bmatrix}2&1\\0&0\end{bmatrix}\begin{bmatrix}0&1\\1&0\end{bmatrix}\begin{bmatrix}0&0\\3&1\end{bmatrix}\begin{bmatrix}0&0\\0&3\end{bmatrix}\right\}$

1. 求从基（I）到基（II）的过渡矩阵C。
2. 求向量 $\begin{bmatrix}7&3\\1&2\end{bmatrix}$ 在基（II）的坐标Y。

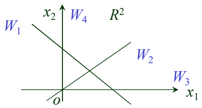
解：可看作空间R^4中的问题，则解法与例11类似。

例题12(P7): 求$P_4[x]$中向量在例10基下的坐标。

§ 1.1　五、子空间 Subspace

- **概述**：线性空间$V_n(F)$中，向量集合V的子集合可以有集合的运算和关系：

  W_1，$W_2 \subseteq V$: $W_1 \cup W_2$，$W_1 \cap W_2$，
- **问题**：这些关系或运算的结果是否仍然为线性空间？空间的分解（子空间表示）？



37

1、子空间的概念

- 定义1.5：设集合$W \subseteq V_n(F)$，$W \neq \varnothing$，

如果W中的元素关于$V_n(F)$中的线性运算也构成线性空间，

则称W是$V_n(F)$的一个子空间。

- 任何线性空间$V_n(F)$，均有两个平凡子空间：

  $V_n(F)$ 和 {0} (零元素空间, 规定维数为0)

38

1、子空间的概念

- 定义1.5：设集合$W \subseteq V_n(F)$，$W \neq \varnothing$，如果W中的元素关于$V_n(F)$中的线性运算也构成线性空间，则称W是$V_n(F)$的一个子空间。
- 判别方法：定理1.5
- W是子空间 ⇔ W对$V_n(F)$的线性运算封闭。
  - 子空间本身就是线性空间。
  - 子空间的判别方法可以作为判别某些线性空间的方法。

39

例 14 R^n，$R^{m\times n}$ 的集合是否为子空间？

- $W_1 = \{ (x_1, x_2) \mid x_1+x_2=1 \}$,

  $W_2 = \{ (x_1, x_2) \mid x_1 - x_2=0 \}$,

  $W_3 = \{ (x_1, 0) \mid x_1$ 任意 $\}$,

  $W_4 = \{ (x_1, x_2, x_3) \mid x_1+2x_2+3x_3=0 \}$,

  $W_5 = \{ (x_1, x_2, x_3) \mid x_1+x_2=0 \}$, ……

W_1不是；$W_2 \sim W_5$是。　$W_2 \sim W_5$的维数？

40

例 14 R^n，$R^{m\times n}$ 的集合是否为子空间？

- $A \in F^{n\times n}$中集合

  $W_1 = \{ A \in F^{n\times n} \mid A^T = A \}$,

  $W_2 = \{ A \in F^{n\times n} \mid A^T = -A \}$,

  $W_3 = \{ A \in F^{n\times n} \mid |A| = 1 \}$。

W_1，W_2是；W_3不是。

W_1，W_2的维数？

41

- **几个重要的子空间（例15，16）：**
- 设向量组$\{ \alpha_1, \alpha_2, \cdots, \alpha_m \} \subseteq V_n(F)$，由它们的所有线性组合生成的子空间：

  $$L\{ \alpha_1, \alpha_2, \cdots, \alpha_m \} = \{ \sum_{i=1}^{m} k_i\alpha_i \mid k_i \in F \}$$
- 矩阵$A \in F^{m\times n}$，两个子空间：
  - A的零空间：$N(A) = \{ X \in F^n : AX=0 \} \subseteq F^n$,
  - A的列空间(值空间)：

    $R(A) = L\{ A_1, A_2, \cdots, A_n \} \subseteq F^m$,

    A_i为A的第i列。

    $R(A) = \{ y : \exists x \in F^n, y = Ax \}$

42

- **几个重要的子空间（例15，16）：**
- 生成子空间：

  $R^n = L\{ e_1, e_2, \cdots, e_n \}$; $L\{ e_1, e_2\}$; …

  $R^{m\times n} = L\{ E_{11}, E_{12}, \cdots, E_{1n} \cdots, E_{m1} \cdots, E_{mn} \}$;

  $L\{ E_{11}, E_{12} \} \subseteq R^{m\times n}$; ……
- $P_n[x] = L\{ 1, x, x^2, \ldots, x^{n-1} \}$

43

2、子空间的运算：“交空间”与“和空间”

- 讨论：设$W_1 \subseteq V_n(F)$，$W_2 \subseteq V_n(F)$，且都是子空间，问$W_1 \cap W_2$和$W_1 \cup W_2$是否仍然是子空间？

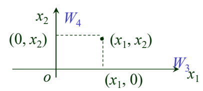
1. （1）交空间
   - 交集：$W_1 \cap W_2 = \{ \alpha \mid \alpha \in W_1$ 而且 $\alpha \in W_2 \} \subseteq V_n(F)$
   - 定理1.6 (1) $W_1 \cap W_2$是子空间，被称为“交空间”

（2）和空间

$W_1 \cup W_2 \neq W_1 + W_2$

- 集合的和集：

  $W_1 + W_2 = \{ \alpha = X_1 + X_2 \mid X_1 \in W_1, X_2 \in W_2 \}$，



$R^2 = W_3 + W_4$

44

2、子空间的运算：“交空间”与“和空间”

- 定理1.6 (2) $W_1 + W_2$是子空间，被称为“和空间”。

$W_1 \cup W_2$一般不是子空间，$W_1 \cup W_2 \subseteq W_1 + W_2$

45

- 例17 设R^3中的子空间$W_1=L\{e_1\}$，$W_2=L\{e_2\}$
  - 求和空间W_1+W_2。
  - 比较：集合$W_1\cup W_2$和集合W_1+W_2。

如果 $W_1=L\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$，
$W_2=L\{\beta_1, \beta_2, \ldots, \beta_k\}$，
则 $W_1+W_2=L\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_k\}$

46

3、维数公式

- 子空间的包含关系：

$$W_1\cap W_2 \subseteq \begin{matrix} W_1 \\ W_2 \end{matrix} \subseteq W_1+W_2 \subseteq V_n(F)$$

$\dim W_1\cap W_2 \le \dim W_i \le \dim W_1+W_2 \le \dim V_n(F)$。

- 定理1.7 （维数定理）

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

证明思路：基扩充方法（从$W_1\cap W_2$的基出发）

47

4、子空间的直和（空间分解）

- 分析 由维数公式知，如果 $\dim(W_1\cap W_2)\neq 0$，则 $\dim(W_1+W_2)<\dim W_1+\dim W_2$

所以，

$\dim(W_1+W_2)=\dim W_1+\dim W_2$
$\Leftrightarrow \dim(W_1\cap W_2)=0$
$\Leftrightarrow W_1\cap W_2=\{0\}$

- 直和的定义：

定义1.6 设$W=W_1+W_2$，若$\dim(W_1\cap W_2)=0$，则称此和为直和，称W为W_1和W_2的直和子空间，记为 $W=W_1\oplus W_2$。

48

- 子空间的"和"为"直和"的充要条件

定理1.8 设 $W=W_1+W_2$，则下列各条等价：

（1） $W=W_1\oplus W_2$
（2） $\forall X\in W$，$X=X_1+X_2$的表示是惟一的
（3） W中零向量的表示是惟一的
（4） $\dim W=\dim W_1+\dim W_2$

证明：循环证法 (1)→(2)→(3)→(4)→(1)

49

- 例1 P12 例18
- 例2 设在$R^{n\times n}$中，子空间

$W_1=\{A \mid A^T=A\}$，$W_2=\{B \mid B^T=-B\}$，

试证：$R^{n\times n}=W_1\oplus W_2$。

证明：(1) 证 $R^{n\times n}=W_1+W_2$
$A=(A+A^T)/2+(A-A^T)/2$
(2) 证 $W_1\cap W_2=\{0\}$
若 $A^T=A$，$A^T=-A$，则 $A=0$

- 例3 子空间W的"直和补子空间"U:

$V_n=W\oplus U$

50

1.2 内积空间

主题：定义内积的概念，借助于内积建立线性空间的度量关系。

一、欧氏空间和酉空间

1 几何空间中度量的定义基础

2 内积的定义

定义1.7 (P13)：要点

1. $(\alpha,\beta)=\overline{(\beta,\alpha)}$
2. $(k\alpha,\beta)=k(\alpha,\beta)$
   $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$
3. $(\alpha,\alpha)\ge 0$;
   $(\alpha,\alpha)=0 \Leftrightarrow \alpha=0$

- 内积(α, β)是二元运算：$V_n(F)\times V_n(F) \longrightarrow F$
- (α, β)的公理性质（对称性，线性性，正定性）
- (α, β)是任何满足定义的运算（映射）。
- 讨论 $(\alpha, \beta_1+\beta_2)$, $(\alpha, k\beta)$, $(\alpha, k_1\beta_1+k_2\beta_2)$, $(0,\alpha)$

$=(\alpha,\beta_1)+(\alpha,\beta_2)$ $\quad =\bar{k}(\alpha,\beta)$ $\quad =\overline{k_1}(\alpha,\beta_1)+\overline{k_2}(\alpha,\beta_2)$ $\quad =0$

51

3 内积空间的定义

- 赋予内积的线性空间称为内积空间，记为 $[V_n(F); (\alpha,\beta)]$；

F＝R，欧氏空间，F＝C，酉空间

4 常见的内积空间：

- $[R^n; (\alpha,\beta)=\alpha^T\beta]$，$(\alpha^T\beta=\beta^T\alpha)$
- $[C^n; (\alpha,\beta)=\beta^H\alpha]$，$\beta^H$为$\beta$的共轭转置，
- $[C^{m\times n}; (A,B)=tr(B^HA)]$
- $[P_n[x]; (f,g)=\int_0^l f(x)g(x)dx]$

52

5 向量的长度

- 定义：$\|\alpha\|=\sqrt{(\alpha,\alpha)}$；单位向量：$\|\alpha\|=1$.
- α,β 的"距离"：$\|\alpha-\beta\|$
- 性质：
  - $\|k\alpha\|=|k|\,\|\alpha\|$;
  - 定理1.9 (Cauchy不等式) 等式成立当且仅当α,β线性相关

  $\forall \alpha, \beta\in[V_n(F); (\alpha,\beta)]$，$|(\alpha,\beta)|\le\|\alpha\|\,\|\beta\|$。
  - $\|\alpha+\beta\|\le\|\alpha\|+\|\beta\|$; $\|\alpha-\beta\|\le\|\alpha\|+\|\beta\|$

6 欧氏空间中向量的夹角

- 定义：$\alpha\neq 0$，$\beta\neq 0$，夹角θ定义为：

$$\cos\theta=\frac{(\alpha,\beta)}{|\alpha||\beta|}$$

α和β正交 $\Leftrightarrow (\alpha,\beta)=0$

53

6 线性空间的内积及其计算与矩阵表示：

- 设$\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$是内积空间$V_n(F)$的基，$\forall\alpha, \beta\in V_n(F)$，则有

$$\alpha=x_1\alpha_1+x_2\alpha_2+\ldots+x_n\alpha_n=(\alpha_1\ \alpha_2 \ldots \alpha_n)X;$$
$$\beta=y_1\alpha_1+y_2\alpha_2+\ldots+y_n\alpha_n=(\alpha_1\ \alpha_2 \ldots \alpha_n)Y$$

$$(\alpha,\beta)=\sum_{i=1}^{n}\sum_{j=1}^{n}x_i\bar{y}_j(\alpha_i,\alpha_j)=Y^HAX,$$

度量矩阵 $a_{ij}=(\alpha_i,\alpha_j)$

度量矩阵A的性质：Hermite 性与正定性

定义内积 $\Leftrightarrow$ 在一个基$\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$下定义内积
$\Leftrightarrow$ 确定一个度量矩阵A 。

54

二、标准正交基

1 正交的向量组：

- 定义：{α_1，α_2，…，α_n} 为正交向量组
  $\Leftrightarrow$ $(\alpha_i, \alpha_j) = 0$，$\forall\, i \neq j$
- 性质(定理1.10) 正交向量组线性无关。

2 标准正交基

- 基 {ε_1，ε_2，…，ε_n} 是标准正交基
  $\Leftrightarrow (\varepsilon_i, \varepsilon_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$
- 标准正交基的优点？想想 R^n！

55

标准正交基的优点:

坐标空间F^n的内积

- 度量矩阵是单位矩阵，即A=I
  $\alpha=(\varepsilon_1\,\varepsilon_2 \ldots \varepsilon_n)X$，$\beta=(\varepsilon_1\,\varepsilon_2 \ldots \varepsilon_n)Y$，
  $(\alpha, \beta) = Y^H X$
- $\alpha = x_1\varepsilon_1 + x_2\varepsilon_2 + \ldots + x_n\varepsilon_n$，$x_i = (\alpha, \varepsilon_i)$
- α和β正交 $\Leftrightarrow$ 其坐标 X和Y正交
- 任何向量的内积将对应其坐标空间中的内积

求标准正交基的步骤:

1. Schmidt 正交化（定理1.11）
2. 标准化
3. 矩阵方法讨论

56

求标准正交基的步骤:

1. Schmidt 正交化（定理1.11）

设 {α_1，α_2，…，α_m} 为内积空间线性无关的向量组，则下列方法产生正交向量组：

$$\beta_1 = \alpha_1,\ \beta_k = \alpha_k - \sum_{i=1}^{k-1} \frac{(\alpha_k, \beta_i)}{(\beta_i, \beta_i)} \beta_i,\ k = 2, \cdots, m$$

改写为 $\alpha_k = \sum_{i=1}^{k-1} \frac{(\alpha_k, \beta_i)}{(\beta_i, \beta_i)} \beta_i + \beta_k,\ k = 1, \cdots, m$

由此推知 $L\{\alpha_1, \alpha_2, \cdots, \alpha_m\} = L\{\beta_1, \beta_2, \cdots, \beta_m\}$

2. 标准化 $\varepsilon_k = \frac{\beta_k}{\|\beta_k\|},\ k = 1,\ 2,\ \cdots,\ m$

合二为一 $\beta_1 = \alpha_1,\ \varepsilon_1 = \frac{\beta_1}{\|\beta_1\|},$

$$\beta_k = \alpha_k - \sum_{i=1}^{k-1} (\alpha_k, \varepsilon_i)\varepsilon_i,\ \varepsilon_k = \frac{\beta_k}{\|\beta_k\|},\ k \geq 2$$

3. 矩阵表示 $\alpha_k = \sum_{i=1}^{k-1} (\alpha_k, \varepsilon_i)\varepsilon_i + \varepsilon_k \|\beta_k\|,\ k = 1, \cdots, m$

57

求标准正交基的步骤: 3. 矩阵表示

$$\alpha_k = \sum_{i=1}^{k-1} \frac{(\alpha_k, \beta_i)}{(\beta_i, \beta_i)} \beta_i + \beta_k = \sum_{i=1}^{k-1} (\alpha_k, \varepsilon_i)\varepsilon_i + \varepsilon_k \|\beta_k\|,\ k = 1, \cdots, m$$

$$(\alpha_1, \alpha_2, \cdots, \alpha_m) = (\beta_1, \beta_2, \cdots, \beta_m) \begin{bmatrix} 1 & \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)} & \cdots & \frac{(\alpha_m, \beta_1)}{(\beta_1, \beta_1)} \\ & 1 & \cdots & \frac{(\alpha_m, \beta_2)}{(\beta_2, \beta_2)} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$

$$(\alpha_1, \alpha_2, \cdots, \alpha_m) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m) \begin{bmatrix} \|\beta_1\| & (\alpha_2, \varepsilon_1) & \cdots & (\alpha_m, \varepsilon_1) \\ & \|\beta_2\| & \cdots & (\alpha_m, \varepsilon_2) \\ & & \ddots & \vdots \\ & & & \|\beta_m\| \end{bmatrix}$$

m=n, 得基变换公式！

58

[R^n；$(\alpha, \beta) = \beta^T\alpha$]，[C^n；$(\alpha, \beta) = \beta^H\alpha$]的标准正交基均为自然基{e_i}；

[$R^{m\times n}$; (A, B) = tr (B^TA)]，[$C^{m\times n}$; (A, B) = tr (B^HA)]的标准正交基均为自然基{E_{ij}}。

"正交补"子空间

(i) 集合的U的正交集:

- $U^\perp$= {$\alpha \in V_n(F)$：$\forall \beta \in U$，$(\alpha, \beta) = 0$ }

(ii) 若U是$V_n(F)$的子空间，则 $U^\perp$ 也是$V_n(F)$子空间，称为U的正交补子空间。

(iii) $V_n(F)=U \oplus U^\perp$。

59

求标准正交基的步骤: 公式推导

$$\alpha_k = \sum_{i=1}^{k-1} \frac{(\alpha_k, \beta_i)}{(\beta_i, \beta_i)} \beta_i + \beta_k = \sum_{i=1}^{k-1} (\alpha_k, \varepsilon_i)\varepsilon_i + \varepsilon_k \|\beta_k\|,\ k = 1, \cdots, m$$

设 $\beta_2 = \alpha_2 + l_{2,1}\beta_1$，$l_{2,1}$ 待定，由 $(\beta_2, \beta_1) = 0$，得

$$(\alpha_2, \beta_1) + l_{2,1}(\beta_1, \beta_1) = 0 \quad \Rightarrow\ l_{2,1} = -\frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)}$$

一般地，设已得到正交向量组 $\beta_1,\ \beta_2,\ \cdots,\ \beta_{k-1}$

令 $\beta_k = \alpha_k + \sum_{i=1}^{k-1} l_{k,i}\beta_i,\ k > 1$

由 $(\beta_k, \beta_j) = 0, 1 \leq j \leq k-1$ 得

$$(\alpha_k, \beta_j) + l_{k,j}(\beta_j, \beta_j) = 0 \quad \Rightarrow\ l_{k,j} = -\frac{(\alpha_k, \beta_j)}{(\beta_j, \beta_j)}, k > 1$$

$$\frac{(\alpha_k, \beta_i)}{(\beta_i, \beta_i)}\beta_i = \frac{(\alpha_k, \beta_i)}{\|\beta_i\|^2}\beta_i = (\alpha_k, \frac{\beta_i}{\|\beta_i\|})\frac{\beta_i}{\|\beta_i\|} = (\alpha_k, \varepsilon_i)\varepsilon_i$$

60

---

§ 1.3 线性变换（函数，映射）

一、线性变换的概念

从一般性的角度给出的定义

1. 定义 1.11 (P.19) $V_n(F)$上的线性变换T

要点:

(i) T是$V_n(F)$上的变换:

T: $V_n(F) \to V_n(F)$, $\alpha \to T(\alpha)$ (象)

(ii) T具有线性性:

$T(\alpha+\beta) = T(\alpha)+T(\beta)$

$T(k\alpha) = kT(\alpha)$

合二为一: $T(k_1\alpha + k_2\beta) = k_1T(\alpha) + k_2T(\beta)$

61

线性变换的例子

例24 $V_n(F)$上的相似变换T_λ：λ是F中给定的数，

$\forall \alpha \in V_n(F)$，$T_\lambda(\alpha) = \lambda\alpha$。

- 特例：$\lambda=1$，T_1是恒等变换: $T_1(\alpha) = \alpha$，
  $\lambda=0$，T_0是零变换: $T_0(\alpha) = 0$。

可以在任何线性空间中 定义相似变换!

例25 $P_n[x]$中的微分变换

(积分变换？线性但不是$P_n[x]$上的)

例26 F^n上的变换 T_A：设 $A \in F^{n\times n}$是一个给定的矩阵，

$\forall\, x \in F^n$，$T_A(x)=Ax$。

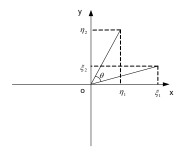
62

例

二维实向量空间 $R^2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \middle| x_i \in R \right\}$，将其绕原点旋转$\theta$角的操作就是一个线性变换。

[证明] $\forall x \in R^2, x = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$ $\quad y = Tx = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}$ $\quad \begin{cases} \eta_1 = x_1\cos\theta + x_2\sin\theta \\ \eta_2 = -x_1\sin\theta + x_2\cos\theta \end{cases}$

$$\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \in R^2$$



63

2 线性变换的性质：

$T(k_1\alpha+k_2\beta)=k_1T(\alpha)+k_2T(\beta)$

(1) $T(0) = 0$

(2) $T(-\alpha) = -T(\alpha)$

(3) $T\sum_{i=1}^{m}k_i\alpha_i = \sum_{i=1}^{m}k_iT(\alpha_i)$

线性变换保持线性相关性不变！

(4) 若$\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$线性相关，则

$\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_m)\}$也线性相关。

3 线性变换的象空间和零空间(~矩阵的列/零空间)

定理1.12 设T为$V_n(F)$上的线性变换，则下述集合均为V_n的子空间（分别称为T的象空间和零空间）：

象空间 $R(T)=\{\beta: \exists\, \alpha\in V_n(F), \beta = T(\alpha)\}$

零空间 $N(T)=\{\alpha: \alpha\in V_n(F), T(\alpha) = 0\}$ (证明)

定义：T 的秩 = dim R(T)； T 的零度 = dim N(T)

例 27 求F^n中的线性变换T_A: $Y=AX$ 的象空间和零空间。

$R(T_A) = R(A)$； $N(T_A) = N(A)$

4 线性变换的"运算"(~ 函数的运算)

设T，T_1，T_2都是$V_n(F)$上的线性变换，它们的下述运算均构成$V_n(F)$上的线性变换：

(i) 加法 T_1+T_2: $\forall\alpha\in V_n(F)$,

$(T_1+T_2)(\alpha) = T_1(\alpha)+T_2(\alpha)$

(ii) 乘法 T_1T_2: $\forall\alpha\in V_n(F)$，$(T_1T_2)(\alpha) = T_1(T_2(\alpha))$

(iii) 数乘 kT: $\forall\alpha\in V_n(F)$，$(kT)(\alpha) = k(T(\alpha))$

(iv) 可逆变换T: $\exists\, T_1$使得, $TT_1 = T_1T = I$, 记$T_1 = T^{-1}$；

$T^{-1}(\beta) = \alpha \Leftrightarrow T(\alpha) = \beta$。

(v) 乘方变换：$T^m = TTTT\ldots T$ (m个T相乘)

注意：变换乘法一般不具有交换律，如同矩阵乘法；

$V_n(F)$上的线性变换的全体Φ构成线性空间Φ(F)！

二、 线性变换的矩阵

$T(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A$

$\Leftrightarrow T(\alpha_i)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A_i$

$V_n(F)$上的变换T $\xleftrightarrow[\text{给定基}]{}$ $F^{n\times n}$中的矩阵A

1 线性变换的矩阵与变换的坐标式

$V_n(F)$上线性变换的特点分析：

- 定义变换T ⇔ 确定基中向量的象$T(\alpha_i)$
- 定义$T(\alpha_i)$ ⇔ 确定它在基下$\{\alpha_i\}$的坐标A_i
- 定义变换T ⇔ 确定矩阵$A = [A_1, A_2, \ldots, A_n]$

(i) A为变换矩阵: $T(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\alpha_1, \alpha_2, \ldots, \alpha_n)A$

(ii) 变换的坐标式: 设 $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)X$,

$T(\alpha) = (\alpha_1, \alpha_2, \ldots, \alpha_n)Y$，则 $Y=AX$ ~线性变换T_A

常见线性变换的矩阵

$T(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A$

$\Leftrightarrow T(\alpha_i)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A_i$

相似变换T_λ下的矩阵：λI

$$T_\lambda(\alpha_i)=\lambda\alpha_i=(\alpha_1,\alpha_2,\cdots,\alpha_n)(\lambda e_i), i=1,2,\cdots,n$$

$$\Rightarrow T_\lambda(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)(\lambda e_1,\lambda e_2,\cdots,\lambda e_n)$$

线性变换T_A在自然基下的矩阵： A

$$T_A(e_i)=Ae_i=A_i=I\cdot A_i=(e_1,e_2,\cdots,e_n)A_i, i=1,2,\cdots,n$$

$$\Rightarrow T_A(e_1,e_2,\cdots,e_n)=(A_1,A_2,\cdots,A_n)=A=IA=(e_1,e_2,\cdots,e_n)A$$

F^n上线性变换T在自然基下的矩阵：$T(e_1\ e_2\ \ldots\ e_n)$

$$T(e_i)=I\cdot T(e_i)=(e_1,e_2,\cdots,e_n)T(e_i), i=1,2,\cdots,n$$

$$\Rightarrow T(e_1,e_2,\cdots,e_n)=I\cdot T(e_1,e_2,\cdots,e_n)$$

常见线性变换的矩阵

$T(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A$

$\Leftrightarrow T(\alpha_i)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A_i$

$P_n[x]$中的微分变换在自然基下的矩阵：

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 \\ \vdots & \vdots & \vdots & \cdots & n-1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$\frac{d}{dx}(x^k)=kx^{k-1}=(1,x,x^2,\cdots,x^{n-1})\begin{pmatrix} 0 \\ \vdots \\ k \\ \vdots \\ 0 \end{pmatrix}$$

$k=0,1,2,\cdots,n-1$

例题 对线性变换：$D=\dfrac{d}{dx}$ $P_4[X] \longrightarrow P_4[X]$，

1 求D在基$\{1, X, X^2, X^3\}$下的变换矩阵。

2 求向量 $p(x)=10-2x+2x^2+3x^3$ 在变换D下的象。$D\sum_{k=0}^{n-1}a_kx^k=\sum_{k=1}^{n-1}ka_kx^{k-1}$

2 线性变换运算的矩阵对应（Th1.13）：

- 设$V_n(F)$上的线性变换T_1，T_2，它们在同一组基下的矩阵：$T_1 \leftrightarrow A_1$；$T_2 \leftrightarrow A_2$

（i） $(T_1+T_2) \leftrightarrow (A_1+A_2)$

（ii） $(T_1T_2) \leftrightarrow A_1A_2$

（iii） $(kT) \leftrightarrow kA$

（iv） $T^{-1} \leftrightarrow A^{-1}$

(i), (iii)合并：$(k_1T_1+k_2T_2) \leftrightarrow (k_1A_1+k_2A_2)$

3 不同基下的变换矩阵

$T(\sum_{i=1}^{n}x_i\alpha_i)=\sum_{i=1}^{n}x_iT(\alpha_i)$

两组基：$\{\alpha_1, \alpha_2, \ldots,\alpha_n\}$，$\{\beta_1, \beta_2, \ldots, \beta_n\}$，

$(\beta_1\,\beta_2\ldots\,\beta_n) = (\alpha_1\,\alpha_2\ldots\,\alpha_n)C$

$T(\alpha_1\,\alpha_2\ldots\,\alpha_n) = (\alpha_1\,\alpha_2\ldots\,\alpha_n)A$

$T(\beta_1\,\beta_2\ldots\,\beta_n) = (\beta_1\,\beta_2\ldots\,\beta_n)B$

$\Rightarrow$ $B=C^{-1}AC$

Th1.14 同一个线性变换在不同基下的矩阵是相似的

$$T((\alpha_1,\alpha_2,\cdots,\alpha_n)X)=(T(\alpha_1,\alpha_2,\cdots,\alpha_n))X$$

$$T((\alpha_1,\alpha_2,\cdots,\alpha_n)C)=(T(\alpha_1,\alpha_2,\cdots,\alpha_n))C$$

$$T(\beta_1,\beta_2,\cdots,\beta_n)=(T(\alpha_1,\alpha_2,\cdots,\alpha_n))C=(\alpha_1,\alpha_2,\cdots,\alpha_n)AC$$

$$(\beta_1,\beta_2,\cdots,\beta_n)B=(\alpha_1,\alpha_2,\cdots,\alpha_n)AC=(\beta_1,\beta_2,\cdots,\beta_n)C^{-1}AC$$

例28(P23) 给定R^3上的线性变换

$T((x_1, x_2, x_3)^T) = (x_1+2x_2+x_3, x_2-x_3, x_1+x_3)^T$,

求T在基$\alpha_1=(1\ 0\ 1)^T$, $\alpha_2=(0\ 1\ 1)^T$, $\alpha_3=(1\ -1\ 1)^T$下的变换矩阵B。

例29(P24)

设单位向量 u =（2/3, –2/3, –1/3），给定R^3上的线性变换 $P(x) = x - (x, u)u$，

1. 求P在自然基$\{e_1, e_2, e_3\}$下的变换矩阵A。
2. 求P在标准正交基$\{u_1, u_2, u_3\}$下的变换矩阵B。（直接按定义；或同前利用Th1.14）

$T(k_1\alpha + k_2\beta) = k_1T(\alpha) + k_2T(\beta)$

$T(\alpha_1,\alpha_2,\cdots,\alpha_n) = (\alpha_1,\alpha_2,\cdots,\alpha_n)A$

$\Leftrightarrow T(\alpha_i) = (\alpha_1,\alpha_2,\cdots,\alpha_n)A_i$

小结

$V_n(F)$上的线性变换T $\xleftrightarrow[\text{给定基}]{}$ $F^{n\times n}$中的矩阵A

线性变换的定义、基本性质及运算；

子空间R(T)和N(T)；

线性变换（在给定基下）的矩阵及变换坐标式；

常见线性变换T及其变换矩阵A；

线性变换T在不同基下的矩阵是相似的。

$$T(L\{\alpha_1,\alpha_2,\cdots,\alpha_m\}) = L\{T(\alpha_1),T(\alpha_2),\cdots,T(\alpha_m)\}$$

$$T(L\{\alpha_1,\alpha_2,\cdots,\alpha_m\}) \ni T(\textstyle\sum_{k=1}^m x_k\alpha_k) = \sum_{k=1}^m x_kT(\alpha_k) \in L\{T(\alpha_1),T(\alpha_2),\cdots,T(\alpha_m)\}$$

73

三、不变子空间

- 问题的背景：
  - 变换矩阵简化和空间分解的对应关系
- 1. 不变子空间的概念
  - 矩阵简化要求空间分解的特点
  - 不变子空间的定义(p24, 定义1.14)

T也是W上的线性变换！

设T是$V_n(F)$上的线性变换，W是V_n的子空间。若∀α ∈W 有 T(α) ∈W (即值域T(W) ⊆W)，则称W是T的不变子空间。

$$T(W) = \{ T(\alpha) \mid \alpha \in W \}$$

74

2 . 不变子空间的判别

W是T的不变子空间 ⟺ ∀α ∈W ⇒ T(α) ∈W。

⟺ T(W) ⊆ W。

- 特别：若W = L{ α_1，α_2，…，α_m }，则

W是T的不变子空间 ⟺ $T(\alpha_i) \in W$。

- P24，例30

R^3上的正交投影P：$P(x) = x - (x, u)u$，u是单位向量。证明L(u)和

$u^\perp = \{ x：(x, u) = 0 \}(= L(u)^\perp)$是P的不变子空间。

P在L(u)上是零变换，在$u^\perp$上是恒等变换！

75

- 3 空间分解与矩阵分解
  - $V_n(F) = W \oplus U$，W，U是T的不变子空间，

$W = L\{ \alpha_1, \ldots, \alpha_r \}$，$U = L\{\alpha_{r+1}, \ldots, \alpha_n \}$

则 T $\xrightarrow{\{\alpha_1, \ldots, \alpha_r, \alpha_{r+1}, \ldots, \alpha_n\}}$ $\begin{bmatrix} A_1 & \\ & A_2 \end{bmatrix}$

$$T(\alpha_i) = (\alpha_1,\alpha_2,\cdots,\alpha_r)A_{1,i}, i = 1,2,\cdots,r$$

$$= (\alpha_1,\alpha_2,\cdots,\alpha_r,\cdots,\alpha_n)\begin{pmatrix} A_{1,i} \\ 0 \end{pmatrix}, i = 1,2,\cdots,r$$

$$T(\alpha_j) = (\alpha_{r+1},\alpha_{r+2},\cdots,\alpha_n)A_{2,j}, j = r+1,r+2,\cdots,n$$

$$= (\alpha_1,\alpha_2,\cdots,\alpha_r,\cdots,\alpha_n)\begin{pmatrix} 0 \\ A_{2,j} \end{pmatrix}, j = r+1,r+2,\cdots,n$$

76

- 3 空间分解与矩阵分解：一般情形
  - $V_n(F) = U_1 \oplus U_2 \oplus \ldots \oplus U_k$，$U_i$是T的不变子空间，

$U_i = L\{\alpha_{i,1}, \ldots, \alpha_{i,n_i}\}, i = 1, 2, \ldots, k$，

则 T ⟶ $A = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{bmatrix}$

矩阵A_i 的阶数 = dim U_i = n_i

特别地，若∀ i, dim(U_i) = 1，则A为对角矩阵！

77

四、正交变换和酉变换(不改变内积的变换)

- 讨论内积空间 [V(F)；(α，β)] 中最重要的一类变换。
- 1 定义1.15 (P25)：(T(α), T(β))=(α, β)
- 2 正交（酉）变换的性质：

定理1.15 T是内积空间V(F)上的线性变换，则下列命题等价：

（1）T是正交(酉)变换；

（2）T保持向量的长度不变；

（3）T把V(F)的标准正交基变成标准正交基； (证(2)→(3))

（4）T在标准正交基下的矩阵是正交(酉)矩阵。

- 3 变换的矩阵：正交矩阵和酉矩阵的性质
  - 正交矩阵C：$C^TC=I$；酉矩阵U： $U^HU=I$
  - 定理1.16(P27) 正交矩阵C和酉矩阵U有如下性质：

(1) |det(C)|=1, |det(U)|=1; (2) $C^{-1}=C^T, U^{-1}=U^H$;

(3) 正交(酉)矩阵的逆、两个正交(酉)矩阵的乘积仍是正交(酉)矩阵；

(4) n阶正交(酉)矩阵的列或行向量组是$R^n(C^n)$中的标准正交基。

78

常见的基本正交变换：

- 平面上的旋转变换T_θ
  - 几何描述：绕坐标原点，逆时针旋转一个θ角。
  - 变换矩阵：在自然基下，$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$
- R^3空间中的旋转变换
  - 几何描述：绕空间中过原点的一条直线L，旋转一个θ角
  - 求变换矩阵的基本思想：寻求空间中的一组特别的基
- R^n空间中的镜像变换 $S(u) = -u$ $R^n = L\{u\} \oplus L(u)^\perp$
  - 定义：$S(x) = x - 2(x, u)u$，u为单位向量。
  - 变换矩阵与几何意义 $A = \begin{bmatrix} -1 & \\ & I_{n-1} \end{bmatrix}$

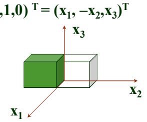
79

- 例题1 设$u = e_2$，镜像变换: $S(x) = x - 2(x, e_2)e_2$

求立方体W在镜像变换下的象。

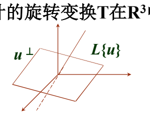
$$S((x_1,x_2,x_3)^T) = (x_1,x_2,x_3)^T - 2x_2(0,1,0)^T = (x_1,-x_2,x_3)^T$$

$$W = \left[\begin{pmatrix}0\\0\\0\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix}\begin{pmatrix}1\\1\\0\end{pmatrix}\begin{pmatrix}0\\1\\0\end{pmatrix}\begin{pmatrix}0\\0\\1\end{pmatrix}\begin{pmatrix}1\\0\\1\end{pmatrix}\begin{pmatrix}1\\1\\1\end{pmatrix}\begin{pmatrix}0\\1\\1\end{pmatrix}\right]$$



- 例题2 求R^3中绕过原点、以 $u = (1, 1, 1)^T$ 为正向的直线，顺u方向看去是逆时针的旋转变换T在R^3中自然基下的变换矩阵。

$R^3 = L\{u\} \oplus L(u)^\perp$ $A = \begin{bmatrix} 1 & \\ & T_\theta \end{bmatrix}$



80

五、线性空间$V_n(F)$ ⟶ $V_m(F)$的线性变换

定义 1.16 (P.28)

- 要点：

(i) ∀ α∈$V_n(F)$，β = T(α) ∈ $V_m(F)$

(ii) T具有线性性：

$$T(\alpha_1 + \alpha_2) = T(\alpha_1) + T(\alpha_2)；T(k\alpha) = kT(\alpha)$$

基本性质

例题1 (P29，例34) $A \in F^{m\times n}$, T_A: F^n ⟶ F^m, $T_A(x)=Ax$

例题2 (P29，例35)

例题3 同构变换σ：$V_n(F)$ ⟶ F^n

例题4 变换T：F^n ⟶ $P_n[x]$

81

$T(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\beta_1,\beta_2,\cdots,\beta_m)A$
$\Leftrightarrow T(\alpha_i)=(\beta_1,\beta_2,\cdots,\beta_m)A_i$

T的变换矩阵：

- T：$V_n(F)\to V_m(F)$　　$A\in F^{m\times n}$
- 设$\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$是空间$V_n(F)$的基，$\{\beta_1, \beta_2, \ldots, \beta_m\}$是空间$V_m(F)$的基，若
  $T(\alpha_1, \alpha_2, \ldots, \alpha_n)=(\beta_1, \beta_2, \ldots, \beta_m)A$
  则称A是T（在给定基下）的变换矩阵。

$R(T)=\{T(\alpha): \alpha\in V_n(F)\}\subseteq V_m(F)$,
$N(T)=\{\alpha: \alpha\in V_n(F), T(\alpha)=0\}\subseteq V_n(F)$。

定理1.17　$\dim R(T)+\dim N(T)=n$。

82

T在不同基下变换矩阵的关系

设在两个空间中分别取两组基：

$$V_n(F):\begin{cases}\{\alpha_1,\alpha_2,\cdots\alpha_n\}\\\{\beta_1,\beta_2,\cdots,\beta_n\}\end{cases}\quad V_m(F):\begin{cases}\{\xi_1,\xi_2,\cdots,\xi_m\}\\\{\eta_1,\eta_2,\cdots,\eta_m\}\end{cases}$$

分析线性变换在两组基下变换矩阵的关系　等价！

$T(\alpha_1, \alpha_2, \ldots, \alpha_n)=(\xi_1, \xi_2, \ldots, \xi_m)A$
$T(\beta_1, \beta_2, \ldots, \beta_n)=(\eta_1, \eta_2, \ldots, \eta_m)B$
$(\beta_1, \beta_2, \ldots, \beta_n)=(\alpha_1, \alpha_2, \ldots, \alpha_n)P$
$(\eta_1, \eta_2, \ldots, \eta_m)=(\xi_1, \xi_2, \ldots, \xi_m)Q$　　$B=Q^{-1}AP$

83

推荐练习题：第一章

P31：

1（3），（4），2，4，6，9，10，13，17，20，23，24，26，28，29，31

84

第2章：Jordan标准形介绍

Jordan Canonical Form

85

第2章 Jordan标准形介绍

- 问题：
  - 对线性空间中的线性变换T，求一组基$\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$和矩阵J，使 T: $\xrightarrow{\{\alpha1,\ \alpha2,\ \ldots,\ \alpha n\}}$ J
    - 简单性：矩阵 J 尽可能简单
    - 通用性：矩阵 J 的结构对任何变换可行
- 思想：
  - 首选 J 为对角形 ⇒ 线性变换的对角化问题。
  - 建立 J 一般的结构 ⇒ Jordan标准形理论。
  - Jordan方法及其应用
- 方法：
  - 矩阵的相似化简问题 ⇒ Jordan化方法
- 重点：

86

$T(\alpha_i)=(\alpha_1,\alpha_2,\cdots,\alpha_n)(\lambda_i e_i)=\lambda_i\alpha_i$

2.1 线性变换的对角表示

- 背景：求基$\{\alpha_i, i=1\sim n\}$, 使得
  $$T(\alpha_1\,\alpha_2\ldots\alpha_n)=(\alpha_1\,\alpha_2\ldots\alpha_n)\begin{bmatrix}\lambda_1&&&\\&\lambda_2&&\\&&\ddots&\\&&&\lambda_n\end{bmatrix}$$

⟺ 1. $\{\alpha_1\,\alpha_2\ldots\alpha_n\}$ 线性无关
2. $L\{\alpha_i\}$是不变子空间: $T\alpha_i=\lambda_i\alpha_i$

- 一、变换T的特征值与特征向量　$(\lambda I-T)(\xi)=O$　$(T-\lambda I)(\xi)=O$

1. 定义2.1 (eigenvalue and eigenvector) $T(\xi)=\lambda\xi$
2. 求解分析(p35 定理2.1)　$T(\xi)=\lambda\xi \sim AX=\lambda X$

> A的特征值就是T的特征值　$(\lambda I-A)X=O$
> A的特征向量是T的特征向量的坐标　$(A-\lambda I)X=O$

87

- 不同基下的矩阵相似(Th1.14)
- 相似矩阵有相同的特征值，与基选择无关，但特征向量一般不同：

设 $AX=\lambda X$，$B=P^{-1}AP$，则有

$PBP^{-1}X=\lambda X$，　即 $B(P^{-1}X)=\lambda(P^{-1}X)$.

T或A的特征值与特征向量的求法：

(1) 选择基及T在此基下的矩阵A；
(2) 求A的特征值：求特征多项式的根$f(\lambda)=0$，其中$f(\lambda)=|\lambda I-A|$，设$\lambda_1,\lambda_2,\ldots,\lambda_n$为全部特征值；
(3) 求A关于λ_i的特征向量：求方程$(\lambda_i I-A)X=0$的非零解X，它是T的特征值对应的特征向量的坐标。

88

- 例1 求$P_n[x]$上微分变换d/dx的特征值与特征向量。

解 分三步：求变换在给定基下的矩阵A；求A的特征值；求A的特征向量。

(1) 自然基下的矩阵
$$A=\begin{pmatrix}0&1&0&\cdots&0\\0&0&2&\cdots&0\\0&0&0&\ddots&0\\\vdots&\vdots&\vdots&&n-1\\0&0&0&\cdots&0\end{pmatrix}$$

(2) 由 $|\lambda I-A|=\lambda^n=0$ 知 $\lambda_1=\lambda_2=\cdots=\lambda_n=0$

(3) 解方程 $(A-0\lambda)X=0$ 得通解

$x_2=x_3=\cdots=x_n=0,\ x_1=k$　即　$X=k(1,0,\cdots,0)^{\mathrm T}$

于是，A关于 $\lambda=0$ 的特征向量为 $X=k(1,0,\cdots,0)^{\mathrm T},k\neq0$,

从而得T=d/dx的特征向量为 $\xi=(1,x,\cdots,x^{n-1})X=k,k\neq0$.

89

例2 设A、B分别为$m\times n$和$n\times m$阶矩阵，证明AB和BA有相同的非零特征值。

证明 $\begin{pmatrix}AB&0\\B&0\end{pmatrix}$ 和 $\begin{pmatrix}0&0\\B&BA\end{pmatrix}$ 相似，则

$$\left|\lambda I_{m+n}-\begin{pmatrix}AB&0\\B&0\end{pmatrix}\right|=\left|\lambda I_{m+n}-\begin{pmatrix}0&0\\B&BA\end{pmatrix}\right|$$

即 $$\begin{vmatrix}\lambda I_m-AB&0\\-B&\lambda I_n\end{vmatrix}=\begin{vmatrix}\lambda I_m&0\\-B&\lambda I_n-BA\end{vmatrix}$$

推出　$\lambda^n|\lambda I_m-AB|=\lambda^m|\lambda I_n-BA|$

因此，AB和BA有相同的非零特征值。

90

3. 特征向量的空间性质

1) 特征子空间：$V_\lambda = \{ \xi \mid T\xi = \lambda\xi \} = N(T-\lambda I)$

2) 特征子空间的性质：(p36，定理2.2)

- ✓ $V_{\lambda i}$是不变子空间
- ✓ $\lambda_i \neq \lambda_j$，则 $V_{\lambda i} \cap V_{\lambda j} = \{0\}$
- ✓ 若λ_i是k_i重特征值，则 $1 \le \dim V_{\lambda i} \le k_i$

推论：

1) 若λ_i是单特征值，则$\dim V_{\lambda i} = 1$

2) $V_{\lambda 1}+V_{\lambda 2}+\ldots+V_{\lambda s} = V_{\lambda 1}\oplus V_{\lambda 2}\oplus\ldots\oplus V_{\lambda s}$

3) $V_{\lambda 1}\oplus V_{\lambda 2}\oplus\ldots\oplus V_{\lambda s} \subseteq V_n(F)$

91

二、线性变换矩阵对角化的充要条件

$$f(\lambda)=\det[\lambda I-A]=(\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k_2}\cdots(\lambda-\lambda_s)^{k_s}$$

$$\sum_{i=1}^{s} k_i = n$$

T可以对角化：存在一组基，使得T在此基下的矩阵是对角阵。这等价于T的变换矩阵可以对角化（因不同基下的矩阵相似）。

定理2.3 T可以对角化 ⇔ T有n个线性无关的特征向量。

⇔ $\Sigma \dim V_{\lambda i} = n$

⇔ $\dim V_{\lambda i} = k_i$, $i=1, \ldots, s$

➡ 定理2.4 T可以对角化 ⇔ $V_{\lambda 1}\oplus V_{\lambda 2}\oplus\ldots\oplus V_{\lambda s} = V_n(F)$

92

例3 $n>1$时，$P_n[x]$上微分变换d/dx没有对角矩阵表示。

例4 幂等矩阵和乘方矩阵的对角表示特性。

- 例题 已知$\{\alpha_1, \alpha_2, \alpha_3\}$是线性空间$V_3(F)$的基，T是$V_3$上如下定义的线性变换，

  $T(\alpha_1) = \alpha_1$

  $T(\alpha_2) = 2\alpha_2$

  $T(\alpha_3) = \alpha_1 + t\,\alpha_2 + 2\,\alpha_3$

  讨论：t 为何值，T 有对角矩阵表示

例题 设 $u=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\end{bmatrix}$，求$R^3$上正交投影$P(x) = x-(x, u)\,u$的特征值和特征向量。

93

2.2 Jordan 矩阵介绍

- 目标：发展一个所有方阵都能与之相似的矩阵结构 --- Jordan矩阵。
- 一、Jordan 矩阵

1. Jordan 块(p40，定义2.3)

   1. 形式： ➢ 特征值λ
   2. 确定因素：➢ 矩阵的阶数

   $$J(\lambda)=\begin{bmatrix}\lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & 1 \\ & & & \lambda\end{bmatrix}$$

   3. Jordan 块矩阵的例子：

   例题1 下列矩阵哪些是Jordan块？

$$[2] \quad \begin{bmatrix}2 & 1\\0 & 2\end{bmatrix} \quad \begin{bmatrix}1 & 1\\0 & 2\end{bmatrix} \quad \begin{bmatrix}4 & 1 & 0\\0 & 4 & 0\\0 & 0 & 4\end{bmatrix} \quad \begin{bmatrix}0 & 1 & 0\\0 & 0 & 1\\0 & 0 & 0\end{bmatrix}$$

94

2 Jordan 矩阵

1) 形式: 由Jordan块构成

$$\begin{bmatrix}J_1(\lambda_1) & & & \\ & J_2(\lambda_2) & & \\ & & \ddots & \\ & & & J_m(\lambda_m)\end{bmatrix}$$

2) Jordan矩阵举例

3) 特点 ✓ 元素的结构

✓ Jordan矩阵是上三角矩阵

✓ 对角矩阵是Jordan 矩阵

3 Jordan 标准形

定理2.5 (存在定理) 在复数域上，每个方阵A都相似于一个Jordan阵J_A。

含义：➢ Jordan 矩阵可以作为相似标准形。

➢ 惟一性：Jordan 子块的集合惟一。

➢ A相似于B ⇔ J_A相似于J_B

95

4 方阵A的Jordan 标准形的求法

- 目标：求可逆矩阵P和Jordan矩阵J_A，使$AP=PJ_A$
- 分析方法：

  在定理 2.5 的基础上逆向分析矩阵J_A和P的构成。
- 求法与步骤：

$$f(\lambda)=|\lambda I-A|=(\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k_2}\cdots(\lambda-\lambda_s)^{k_s}$$

➢ 矩阵A和J_A的特征值相等 ➡

$$J_A=\begin{bmatrix}J_1(\lambda_1) & & & \\ & J_2(\lambda_2) & & \\ & & \ddots & \\ & & & J_s(\lambda_s)\end{bmatrix}$$

$$AP_i = P_i J_i(\lambda_i)$$

$$J_i(\lambda_i)=diag\{J_{i1}(\lambda_i), J_{i2}(\lambda_i), \cdots, J_{it_i}(\lambda_i)\}, i=1, 2, \cdots, s$$

为k_i阶Jordan阵。$J_{ij}(\lambda_i), j=1, 2, \cdots, t_i$ 为n_{ij}阶Jordan块。

96

➢ 再细分矩阵P_i和J_i，在Jordan块上，有

$$AP_{ij} = P_{ij}J_{ij}(\lambda_i), j=1,2,\cdots,t_i$$

- Jordan链条$P_{ij}=\{\alpha, y_2, \ldots, y_{nj}\}$，确定$P_{ij}$及其列数，即Jordan块$J_{ij}$的阶数$n_j$

$$\begin{cases}(A-\lambda_i I)\alpha = 0 \\ (A-\lambda_i I)y_2 = \alpha \\ (A-\lambda_i I)y_3 = y_2 \\ \cdots \qquad \cdots \\ (A-\lambda_i I)y_{n_j} = y_{n_{j-1}}\end{cases}$$

特征向量

广义特征向量

97

- Jordan标准型的计算步骤（Jordan化方法）：

➢求A的特征值，由特征值λ_i的代数重数k_i确定主对角线元素是的 λ_i 的 Jordan 矩阵$J(\lambda_i)$的阶数；

➢解方程$(A-\lambda_i I)X = 0$，求A关于λ_i的线性无关特征向量（解空间的基），由特征值λ_i对应的线性无关的特征向量的个数t_i（即几何重数$\dim V_{\lambda i}$）确定 $J(\lambda_i)$ 中 Jordan 块的个数；

➢由特征向量求得的Jordan 链条的长度确定Jordan块的阶数；

➢链条中的向量合起来构成可逆矩阵P，Jordan块构成J_A。

例题1, 2 (p44，例题5；p45，例题6) 给定A，求可逆阵P和J_A使 $P^{-1}AP = J_A$。

98

- 例题3 将矩阵A化为Jordan 矩阵。

$$A=\begin{bmatrix}3 & 4 & 0 & 0\\-1 & -1 & 0 & 0\\0 & 0 & 2 & 1\\0 & 0 & -1 & 0\end{bmatrix} \quad J_A=\begin{bmatrix}1 & 1 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & 1 & 1\\0 & 0 & 0 & 1\end{bmatrix} \quad or \quad \begin{bmatrix}1 & 0 & 0 & 0\\0 & 1 & 1 & 0\\0 & 0 & 1 & 1\\0 & 0 & 0 & 1\end{bmatrix}$$

解 1. $|\lambda I-A|=(\lambda-1)^4=0$, 得四重根 $\lambda=1$.

2. 解方程 $(I-A)X=0$, 得通解

$$X=l_1(-2,1,0,0)^T+l_2(0,0,-1,1)^T.$$

知有两个Jordan块！$\Leftarrow t=4-r(I-A)=2$

$\alpha_1=(-2,1,0,0)^T \Rightarrow \beta_1=(-1,0,0,0)^T$; (可推知$J_A$)！

$\alpha_2=(0,0,-1,1)^T \Rightarrow \beta_2=(0,0,-1,0)^T, \Rightarrow P=(\alpha_1,\beta_1,\alpha_2,\beta_2)$.

99

例题4 (p46，例题7) 设$P_3[x]$上线性变换T在自然基下的矩阵为A，求$P_3[x]$的基使得T在此基下的矩阵为Jordan矩阵。其中

$$A=\begin{bmatrix}2&-1&-1\\2&-1&-2\\-1&1&2\end{bmatrix}.$$

解 分析：因$P^{-1}AP=J_A$，故由Th1.14知，P为自然基到待求基的过渡矩阵。求得P，便可得到所求！

$$|\lambda I-A|=(\lambda-1)^3=0;t=3-r(I-A)=2\quad\Rightarrow J_A=\begin{bmatrix}1&0&0\\0&1&1\\0&0&1\end{bmatrix}.$$

$(I-A)X=0$ 的通解：$X=l_1(1,1,0)^T+l_2(1,0,1)^T$.

此例，分别以两个特解出发均无解！

故而需以通解代入，再求得一个广义特征值。

100

例题5(p47，例题8) 设A为阶方阵，证明矩阵A和A^T相似。

证明思想：

证明A和A^T相似

$\Leftrightarrow$ 证明 Jordan 矩阵J_A和J_A^T相似，

$\Leftrightarrow$ 证明 J_A和J_A^T的Jordan 块J和J^T相似。

证明方法：

取逆向（反）单位矩阵S，$S=\begin{bmatrix}&&&1\\&&1&\\&\iddots&&\\1&&&\end{bmatrix}$

证明：$S^{-1}=S$，$SJS=J^T$

（backward identity）

101

§ 2.3 最小多项式 (minimal polynomials)

- 讨论 *n* 阶矩阵多项式的相关问题：
  - 矩阵多项式（重点是计算）
  - 矩阵的化零多项式（Cayley 定理）
  - 最小多项式
- Jordan标准形的应用（简化计算）
  - 相似不变性
  - Jordan化的方法

102

一、矩阵多项式

1. 定义 $g(\lambda)=a_m\lambda^m+a_{m-1}\lambda^{m-1}+\cdots+a_1\lambda+a_0$

$$g(A)=a_mA^m+a_{m-1}A^{m-1}+\cdots+a_1A+a_0I$$

2 . 性质（定理2.6）

- $AX=\lambda_0X\ \Rightarrow\ g(A)X=g(\lambda_0)X$
- $P^{-1}AP=B\ \Rightarrow\ P^{-1}g(A)P=g(B)$
- $A=\begin{bmatrix}A_1&&&\\&A_2&&\\&&\ddots&\\&&&A_k\end{bmatrix}\Rightarrow g(A)=\begin{bmatrix}g(A_1)&&&\\&g(A_2)&&\\&&\ddots&\\&&&g(A_k)\end{bmatrix}$

103

3 矩阵多项式$g(A)$的计算

Jordan块

$$A=P\begin{bmatrix}J_1(\lambda_1)&&&\\&J_2(\lambda_2)&&\\&&\ddots&\\&&&J_k(\lambda_k)\end{bmatrix}_{n\times n}P^{-1}\longrightarrow g(A)=P\begin{bmatrix}g(J_1)&&&\\&g(J_2)&&\\&&\ddots&\\&&&g(J_k)\end{bmatrix}_{n\times n}P^{-1}$$

$$J(\lambda)=\begin{bmatrix}\lambda&1&&\\&\lambda&1&\\&&\ddots&1\\&&&\lambda\end{bmatrix}_{r\times r}\longrightarrow g(J)=\begin{bmatrix}g(\lambda)&g'(\lambda)&\frac{g''(\lambda)}{2!}&\cdots&\frac{g^{(r-1)}(\lambda)}{(r-1)!}\\&g(\lambda)&g'(\lambda)&\ddots&\vdots\\&&g(\lambda)&\ddots&\frac{g''(\lambda)}{2!}\\&&&\ddots&g'(\lambda)\\&&&&g(\lambda)\end{bmatrix}$$

$m\geq r$

$g(J)$ 的结构特点： 由第一行的元素生成

104

3 矩阵多项式$g(A)$的计算

$$C_k^i=\frac{k!}{i!(k-i)!}\qquad J(\lambda)=\begin{bmatrix}\lambda&1&&\\&\lambda&1&\\&&\ddots&1\\&&&\lambda\end{bmatrix}_{r\times r}=\lambda I_r+U_r$$

$$J^k=(\lambda I_r+U_r)^k=\sum_{i=0}^kC_k^i\lambda^{k-i}U^i$$

$$\begin{aligned}g(J)&=\sum_{k=0}^ma_kJ^k=\sum_{k=0}^m\sum_{i=0}^ka_kC_k^i\lambda^{k-i}U^i\\&=\sum_{i=0}^m\Big(\sum_{k=i}^ma_kC_k^i\lambda^{k-i}\Big)U^i\\&=\sum_{i=0}^m\frac{1}{i!}\Big(\sum_{k=i}^ma_k\frac{k!}{(k-i)!}\lambda^{k-i}\Big)U^i\\&=\sum_{i=0}^m\frac{1}{i!}\Big(\frac{d^i}{d\lambda^i}\sum_{k=i}^ma_k\lambda^k\Big)U^i=\sum_{i=0}^m\frac{1}{i!}g^{(i)}(\lambda)U^i\end{aligned}$$

105

3 矩阵多项式$g(A)$的计算

$$g(J)=\sum_{k=0}^ma_kJ^k=\sum_{i=0}^m\frac{1}{i!}g^{(i)}(\lambda)U^i$$

$$U^i=\begin{bmatrix}0&\cdots&0&\overset{i+1}{1}&0&\cdots&0\\&0&\cdots&0&1&\ddots&\vdots\\&&\ddots&\ddots&\ddots&\ddots&0\\&&&\ddots&\ddots&\ddots&1\\&&&&\ddots&\ddots&0\\&&&&&\ddots&\vdots\\&&&&&&0\end{bmatrix}_{r\times r}\qquad g(J)=\begin{bmatrix}g(\lambda)&g'(\lambda)&\frac{g''(\lambda)}{2!}&\cdots&\frac{g^{(r-1)}(\lambda)}{(r-1)!}\\&g(\lambda)&g'(\lambda)&\ddots&\vdots\\&&g(\lambda)&\ddots&\frac{g''(\lambda)}{2!}\\&&&\ddots&g'(\lambda)\\&&&&g(\lambda)\end{bmatrix}$$

106

例题1 设 $g(\lambda)=\lambda^3+4\lambda^2-5\lambda+1$

对P44，例5中的矩阵A，计算$g(A)$。

解

$$A=\begin{bmatrix}-3&3&-2\\-7&6&-3\\1&-1&2\end{bmatrix}=P\begin{bmatrix}1&&\\&2&1\\&&2\end{bmatrix}P^{-1}$$

$$g(A)=P\begin{bmatrix}g(1)&&\\&g(2)&g'(2)\\&&g(2)\end{bmatrix}P^{-1}=P\begin{bmatrix}1&&\\&15&23\\&&15\end{bmatrix}P^{-1}$$

代入P可得所求。

107

二、矩阵的化零多项式
(Annihilating polynomials of Matrices)

- 问题：设$A\in F^{n\times n}$，$A\neq 0$，问是否存在非零多项式$g(\lambda)$，使得$g(A)=0$？

1. 化零多项式（*P.52*）

如果 $g(A)=0$，则称$g(\lambda)$为矩阵A的化零多项式。

要点：若A有化零多项式，则有无穷多化零多项式；$g(A)=0$ 的决定因素和存在性问题。

- ***Cayley-Hamilton*** 定理（*P.52*， *定理2.7*）：

设 $A\in F^{n\times n}$，$f(\lambda)=\det(\lambda I-A)$，则$f(A)=0$。

证明：Jordan化方法推知，对任意Jordan块均有 $f(J_i)=0$，从而有 $f(A)=0$。

108

二、矩阵的化零多项式

(Annihilating polynomials of Matrices)

$g(\lambda)=q(\lambda)f(\lambda)+r(\lambda)$

- *Cayley* 定理的应用举例：
  - 使A^k($\forall\ k\geq n$)降阶至不超过n-1次的多项式(除法余项)。
  - 由$f(A)=0$, 知A的逆矩阵可以用多项式表示。
  - 对线性变换T，$f(T)=0$，即$f(T)$为零变换。

$$f(A)=A^n+a_{n-1}A^{n-1}+\cdots+a_1A+a_0I=0$$
$$\Rightarrow A^n=-(a_{n-1}A^{n-1}+\cdots+a_1A+a_0I)$$
$$a_0\neq 0\Rightarrow A(A^{n-1}+a_{n-1}A^{n-2}+\cdots+a_1I)=-a_0I$$
$$A^{-1}=-\frac{1}{a_0}(A^{n-1}+a_{n-1}A^{n-2}+\cdots+a_2A+a_1I)=0$$

109

三、最小多项式

1 定义(P.54，定义2.5)

- $m_A(\lambda)$是最小多项式
  - $m_A(\lambda)=0$
  - $m_A(\lambda)$在化零多项式中次数最低
  - $m_A(\lambda)$最高次项系数是1
  - $\Rightarrow m_A(\lambda)$整除任何化零多项式

2 $m_A(\lambda)$的结构：

设$f(\lambda)=|\lambda I-A|=(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_s)^{r_s}$

- 定理2.8 $m_A(\lambda)=(\lambda-\lambda_1)^{t_1}(\lambda-\lambda_2)^{t_2}\cdots(\lambda-\lambda_s)^{t_s}$

$f(\lambda)$与$m_A(\lambda)$谱相同　　$1\leq t_i\leq r_i$

- 定理2.9 $m_A(\lambda)=(\lambda-\lambda_1)^{\bar{n}_1}(\lambda-\lambda_2)^{\bar{n}_2}\cdots(\lambda-\lambda_s)^{\bar{n}_s}$

$\bar{n}_i$ 是λ_i对应的Jordan块的指数(最高阶数)。

110

3 线性变换有对角矩阵表示的条件

- 讨论线性变换的最小多项式

定理2.10：线性变换T可以对角化的充要条件是T的最小多项式是一次因子的乘积。

推论：若A有一个化零多项式由一次因子构成，则A可对角化。

不可对角化！

例题1 设$A\in R^{4\times 4}$，$m_A(\lambda)=(\lambda-1)(\lambda-2)^2$

求矩阵A的所有可能的Jordan矩阵。$\bar{n}_1=1,\ \bar{n}_2=2$

例题2 设 $g(\lambda)=(\lambda-1)(\lambda-2)(\lambda-4)$

是矩阵A的化零多项式，证明：A相似于对角矩阵。

因$m_A(\lambda)$整除$g(\lambda)$，故$m_A(\lambda)$的因子均为一次！得证。

111

3 线性变换有对角矩阵表示的条件

- 讨论线性变换的最小多项式

例题3（P.56，例10）求$m_A(\lambda)$。　*A不可对角化！*

解　$f(\lambda)=|\lambda I-A|=(\lambda-1)(\lambda-2)^2$

$$\Rightarrow m_T(\lambda)=(\lambda-1)(\lambda-2)\ \ or\ \ (\lambda-1)(\lambda-2)^2$$
$$(A-I)(A-2I)\neq 0\ \Rightarrow\ m_T(\lambda)=(\lambda-1)(\lambda-2)^2$$

例题4（P.56，例11）求$m_A(\lambda)$。　*A不可对角化！*

解　$|\lambda I-A|=(\lambda-1)(\lambda+1)(\lambda-2)^2$　$\Rightarrow\bar{n}_1=\bar{n}_2=1$

$$n-r(A-2I)=4-3=1\quad\Rightarrow\bar{n}_3=2$$
$$\Rightarrow m_T(\lambda)=(\lambda-1)(\lambda+1)(\lambda-2)^2$$

112

矩阵相似问题中的一些结果

$A,B\in R^{n\times n},P^{-1}AP=B.\quad P^{-1}g(A)P=g(B).$

1. 相似矩阵具有：

- 相同的特征值和特征多项式；
- 相同的化零多项式和最小多项式
- 相同的行列式、迹和秩；

$$|A|=\lambda_1\lambda_2\cdots\lambda_n;\ tr(A)=\lambda_1+\lambda_2+\cdots+\lambda_n.$$
$$|\lambda I-A|=\prod_{i=1}^{n}(\lambda-a_{ii})+g(\lambda)\ \ (g\text{的阶}m\leq n-2)$$
$$=\lambda^n-(a_{11}+\cdots+a_{nn})\lambda^{n-1}+a_{n-2}\lambda^{n-2}+\cdots+a_0$$
$$\prod_{i=1}^{n}(\lambda-\lambda_i)=\lambda^n-(\lambda_1+\cdots+\lambda_n)\lambda^{n-1}+\cdots+(-1)^n\prod_{i=1}^{n}\lambda_i$$

113

矩阵相似问题中的一些结果

$\sim g(\lambda)=\lambda^k$

充分性 $J_A^k=0$

2. 幂等矩阵、幂零矩阵和乘方矩阵

- 幂等矩阵（idempotent）：$A^2=A$　$\sim g(\lambda)=\lambda^2-\lambda$
- 幂零矩阵（nilpotent）：$A\neq 0$，k为正整数，$A^k=0$
- 乘方矩阵（involutary）：$A^2=I$　$\sim g(\lambda)=\lambda^2-1$

- A为幂等矩阵的充要条件是A相似于矩阵 $\begin{bmatrix} I_r & \\ & 0\end{bmatrix}$
- A为幂零矩阵的充要条件是A的特征值都是零。
- A为乘方矩阵的充要条件是A相似于矩阵 $\begin{bmatrix} I & \\ & -I\end{bmatrix}$

114

3、矩阵A， A^T， A^H 和A^HA

设A为n阶方阵，则下列结果成立：

1. 矩阵A相似于矩阵A^T(例8)
2. 矩阵A相似于矩阵A^H的充要条件是矩阵的非实数特征值对应的Jordan 块以共轭对出现。　$|\lambda I-A|=0\ \Leftrightarrow\ |\bar{\lambda}I-A^H|=0$

   A相似于A^H的充要条件是J_A相似于J_A^H
3. 矩阵A^HA相似于矩阵AA^H

特征多项式、值相同，并且可以对角化（Th3.10）。

115

4．设矩阵$A\in F^{m\times n}$，矩阵$B\in F^{n\times m}$，则AB和BA的非零特征值相同(例2)。

讨论：若A、B都是n阶方阵，
1. AB和BA的特征多项式是否相同？
2. AB和BA的最小多项式是否相同？
3. AB和BA是否相似？

若A、B都是n阶方阵，则　$A^{-1}(AB)A=BA$

AB和BA: 特征值相同; 特征多项式相同(例2)。

但不一定相似！若A或B可逆，则AB和BA相似！

AB和BA: 行列式和迹均相同！秩不一定相同。

$$A=\begin{pmatrix}1&0\\0&0\end{pmatrix},B=\begin{pmatrix}0&1\\0&0\end{pmatrix}\quad\Rightarrow AB=B,BA=0.$$

116

第二章的推荐练习题

1，2，3，6，8，9，12，13，16，19，20

117

第3章　矩阵的分解

Matrix Factorization and Decomposition

118

矩阵分解的概述

- 矩阵的分解：两种常见的形式
  - $A=A_1+A_2+\ldots+A_k$　　矩阵的和
  - $A=A_1A_2\ldots A_m$　　矩阵的乘积，如FFT
- 矩阵分解的原则与意义：
  - 实际应用的需要。
  - 显示原矩阵的某些特性
  - 矩阵化简的方法与矩阵技术

理论上的需要
计算上的需要

- 主要技巧：
  - 各种标准形的理论和计算方法
  - 矩阵的分块运算和初等变换

119

§ 3.1 常见的矩阵标准形与分解

- 常见的标准形
  - 等价标准形　$A_{m\times n}=P_{m\times m}\begin{bmatrix} I_r & 0 \\ 0 & 0\end{bmatrix}Q_{n\times n}$
  - 相似标准形　$A_{n\times n}=PJ_AP^{-1}$
  - 合同标准形　$A_{n\times n}=CAC^T$　$A^T=A$
- 本节分解：
  - 三角分解　　等价标准形
  - 满秩分解
  - 可对角化矩阵的谱分解　　相似标准形

120

一、矩阵的三角分解（triangular decomposition)

- 方阵的LU和LDV分解（*P.61*）～解方程
  - LU分解：$A\in F^{n\times n}$， 有下三角形矩阵L，上三角形矩阵U，使得 A = LU。
  - LDV分解：$A\in F^{n\times n}$， L、V分别是主对角线元素为1的下三角形和上三角形矩阵，D为对角矩阵，使得 A = LDV。　由LDV可得LU
  - 已知的方法：Gauss-消元法
  - 基本性质
  - LU　$|A|=|LU|=\prod_{i=1}^{n} l_{ii}u_{ii}.\quad |A_k|=|L_kU_k|=\prod_{i=1}^{k} l_{ii}u_{ii}$
  - LDV　$|A|=|LDV|=|D|=\prod_{i=1}^{n} d_i.\quad |L_kD_kV_k|=\prod_{i=1}^{k} d_i$

$$A_n=L_nU_n=\begin{bmatrix} L_k & O \\ * & *\end{bmatrix}\begin{bmatrix} U_k & * \\ O & *\end{bmatrix}=\begin{bmatrix} L_kU_k & * \\ * & *\end{bmatrix},\ A_k=L_kU_k, 1\le k\le n$$

121

一、矩阵的三角分解（triangular decomposition)

- 方阵的LU和LDV分解（*P.61*）～解方程
  - 例题1（P.61eg1）设 $A=\begin{bmatrix} 2 & 2 & 3 \\ 4 & 7 & 7 \\ -2 & 4 & 5\end{bmatrix}$

求A的LU和LDV分解。

$$(A\,|\,I)=\begin{bmatrix} 2 & 2 & 3 & 1 & 0 & 0 \\ 4 & 7 & 7 & 0 & 1 & 0 \\ -2 & 4 & 5 & 0 & 0 & 1\end{bmatrix}\xrightarrow[r_3+r_1]{r_2-2r_1}\begin{bmatrix} 2 & 2 & 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & -2 & 1 & 0 \\ 0 & 6 & 8 & 1 & 0 & 1\end{bmatrix}\xrightarrow{r_3-2r_2}\begin{bmatrix} 2 & 2 & 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & -2 & 1 & 0 \\ 0 & 0 & 6 & 5 & -2 & 1\end{bmatrix}$$

$$U=\begin{bmatrix} 2 & 2 & 3 \\ 0 & 3 & 1 \\ 0 & 0 & 6\end{bmatrix}, P=\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -2 & 1\end{bmatrix}, L=P^{-1}=\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1\end{bmatrix}\Rightarrow A=LU$$

结论：如果矩阵A能用两行互换以外的初等行变换化为阶梯形（上三角阵），则A有LU分解。

122

一、矩阵的三角分解（triangular decomposition)

- 方阵的LU和LDV分解（*P.61*）～解方程
  - 例题1（P.61eg1）设 $A=\begin{bmatrix} 2 & 2 & 3 \\ 4 & 7 & 7 \\ -2 & 4 & 5\end{bmatrix}$

求A的LU和LDV分解。

$$A=LU=\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1\end{bmatrix}\begin{bmatrix} 2 & 2 & 3 \\ 0 & 3 & 1 \\ 0 & 0 & 6\end{bmatrix}$$

进一步，可得LDV分解：

$$A=\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1\end{bmatrix}\begin{bmatrix} 2 & 2 & 3 \\ 0 & 3 & 1 \\ 0 & 0 & 6\end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1\end{bmatrix}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6\end{bmatrix}\begin{bmatrix} 1 & 1 & 3/2 \\ 0 & 1 & 1/3 \\ 0 & 0 & 1/6\end{bmatrix}$$

123

- 三角分解的存在性和惟一性
  - 定理3.1（P.62）：
    - 矩阵的k阶顺主子式：取矩阵的前k行、前k列得到的行列式，k=1，2， …， n。
    - 定理: $A\in F^{n\times n}$有惟一LDV分解的充要条件是A的顺主子式$|A_k|$非零，k =1，2，…，n-1，$|A_0|$=1,

    其中 $D = \mathrm{diag}(d_1,d_2,\ldots,d_n)$, $d_k = |A_k| / |A_{k-1}|$, k=1, …, n。

- 讨论 （1） LDV分解的存在 ⇒ LU分解存在
  （2）矩阵可逆与顺序主子式非零的关系

- 定理3.2（P.64）设矩阵$A\in F^{n\times n}$，rank(A) = k（≤n），如果A的前k阶顺序主子式均非0，则A有LU分解。
- 考虑：LDV分解与LU分解的关系。
- 例题2（P.65 例2）　LY=b, UX=Y
- LU分解的应用举例（例3）：求解线性方程组AX=b。

124

二、矩阵的满秩分解

- 定义3.2 （*P.66* ）　行满秩

列满秩

对秩为r 的矩阵$A\in F^{m\times n}$，若存在秩为r的矩阵$B\in F^{m\times r}$，$C\in F^{r\times n}$，使得A=BC，则称此式为A的满秩分解。

- 定理3.2：任何非零矩阵$A\in F^{m\times n}$都有满秩分解。

证明：价标准型求法（行列变换）

设通过行及列的初等变换把A变为等价标准型，即存在可逆矩阵P，Q，使得

$$A_{m\times n}=P_{m\times m}\begin{bmatrix} I_r & 0 \\ 0 & 0\end{bmatrix}Q_{n\times n}$$

令

$$P=(B,B_2), Q=\begin{pmatrix} C \\ C_2\end{pmatrix}\ \text{得}\ A=(B,O)\begin{pmatrix} C \\ C_2\end{pmatrix}=BC$$

B，C满足要求。证毕。

125

二、矩阵的满秩分解

满秩分解的求法：初等变换

- 方法1：等价标准型求法（行列变换）：求两个逆矩阵！
- 方法2：阶梯型求法（行变换）：只求一个逆矩阵！

例题1（P.68，eg4）

- 方法3：求列的极大无关组及表示（行变换）：不用求逆

例题2 （P.69，eg5）例题3（P.70，eg6）

法2

$$(A\ |\ I)\xrightarrow{\text{行初等变换}}\begin{pmatrix} C & P \\ O & \end{pmatrix}^{r\text{行}},\quad \mathrm{rank}(C)=r=\mathrm{rank}(A)$$

$$PA=\begin{pmatrix} C \\ \mathrm{O}\end{pmatrix}\ \Rightarrow A=P^{-1}\begin{pmatrix} C \\ \mathrm{O}\end{pmatrix}=(B,B_2)\begin{pmatrix} C \\ \mathrm{O}\end{pmatrix}=BC$$

126

二、矩阵的满秩分解

满秩分解的求法：初等变换

- 方法1：等价标准型求法（行列变换）：求两个逆矩阵！
- 方法2：阶梯型求法（行变换）：只求一个逆矩阵！
- 方法3：求列的极大无关组及表示（行变换）：不用求逆

法3

$$A \xrightarrow{\text{行初等变换}} \begin{pmatrix} I_r & S \\ O & O \end{pmatrix}, PA=\begin{pmatrix} I_r & S \\ O & O \end{pmatrix}, \Rightarrow A=P^{-1}\begin{pmatrix} I_r & S \\ O & O \end{pmatrix},$$

$$A=(B,B_2)\begin{pmatrix} I_r & S \\ O & O \end{pmatrix}=(B,BS)=B(I_r,S)\ =BC$$

$$B=?? \qquad A=(A_1,A_2)\Rightarrow B=A_1$$

127

二、矩阵的满秩分解

$$A=\begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ 1 & 0 & 1 \end{bmatrix}$$

满秩分解的求法：初等变换

例题1-2（P.68-69，例4-5,）法2，法3：求A的满秩分解

$$(A\,|\,I)=\begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{r_3-r_1} \begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 0 & -1 & -1 & -1 & 0 & 1 \end{bmatrix} \xrightarrow{r_3+\frac{1}{2}r_2} \begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1/2 & 1 \end{bmatrix}$$

$$C=\begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \end{bmatrix}, P=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 1/2 & 1 \end{bmatrix}, P^{-1}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -1/2 & 1 \end{bmatrix}, B=\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1/2 \end{bmatrix}$$

$$\Rightarrow A=BC=\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1/2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \end{bmatrix}$$

与书上的不同！
少了一步计算。

128

二、矩阵的满秩分解

$$A=\begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ 1 & 0 & 1 \end{bmatrix}$$

满秩分解的求法：初等变换

例题1-2（P.68-69，例4-5,）法2，法3：求A的满秩分解

$$(A\,|\,I)=\begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{r_3-r_1} \begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 0 & -1 & -1 & -1 & 0 & 1 \end{bmatrix} \xrightarrow{r_3+\frac{1}{2}r_2} \begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1/2 & 1 \end{bmatrix}$$

$$C=\begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \end{bmatrix}, P=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 1/2 & 1 \end{bmatrix}, P^{-1}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -1/2 & 1 \end{bmatrix}, B=\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1/2 \end{bmatrix}$$

$$\Rightarrow A=BC=\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1/2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}$$

129

二、矩阵的满秩分解

$$A=\begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ 1 & 0 & 1 \end{bmatrix}$$

满秩分解的求法：初等变换

例题1-2（P.68-69，例4-5,）法2，法3：求A的满秩分解

$$(A\,|\,I)=\begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{r_3-r_1} \begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 0 & -1 & -1 & -1 & 0 & 1 \end{bmatrix} \xrightarrow{r_3+\frac{1}{2}r_2} \begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1/2 & 1 \end{bmatrix}$$

法3

$$A=\begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ 1 & 0 & 1 \end{bmatrix} \xrightarrow[r_3+\frac{1}{2}r_2]{r_3-r_1} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow[r_1-r_2]{\frac{1}{2}r_2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow B=A_1=\begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix}, C=\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad \Rightarrow A=BC=\begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

130

二、矩阵的满秩分解

$$A=\begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ 1 & 0 & 1 \end{bmatrix}$$

满秩分解的求法：初等变换

例题1-2（P.68-69，例4-5,）法2，法3：求A的满秩分解

$$(A\,|\,I)=\begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{r_3-r_1} \begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 0 & -1 & -1 & -1 & 0 & 1 \end{bmatrix} \xrightarrow{r_3+\frac{1}{2}r_2} \begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1/2 & 1 \end{bmatrix}$$

法3

$$A=\begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ 1 & 0 & 1 \end{bmatrix} \xrightarrow[r_3+\frac{1}{2}r_2]{r_3-r_1} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow[\frac{1}{2}r_2]{r_1-r_2} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow B=\begin{bmatrix} 1 & 2 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, C=\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \Rightarrow A=BC=\begin{bmatrix} 1 & 2 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

131

三、可对角化矩阵的谱分解

- 将方阵分解成用谱加权的矩阵和
  - 谱：设 $|\lambda I-A|=(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_s)^{r_s}$

则称 $\{\lambda_1, \lambda_2, \ldots, \lambda_s\}$为矩阵A的谱 。

1. 可对角矩阵的谱分解

- 分析：

$$P^{-1}AP=\begin{bmatrix} \lambda_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \lambda_1 & & & & & & \\ & & & \lambda_2 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \lambda_2 & & & \\ & & & & & & \ddots & & \\ & & & & & & & \lambda_s & \\ & & & & & & & & \ddots \\ & & & & & & & & & \lambda_s \end{bmatrix}=\begin{bmatrix} \lambda_1 I_{r_1} & & & \\ & \lambda_2 I_{r_2} & & \\ & & \ddots & \\ & & & \lambda_s I_{r_s} \end{bmatrix}$$

132

三、可对角化矩阵的谱分解

$$P^{-1}AP=\begin{bmatrix} \lambda_1 I_{r_1} & & & \\ & \lambda_2 I_{r_2} & & \\ & & \ddots & \\ & & & \lambda_s I_{r_s} \end{bmatrix}$$

1. 可对角矩阵的谱分解

- 分析：$P^{-1}AP=$

$$\lambda_1\begin{bmatrix} I_{r_1} & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}+\lambda_2\begin{bmatrix} 0 & & & \\ & I_{r_2} & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}+\cdots+\lambda_s\begin{bmatrix} 0 & & & \\ & \ddots & & \\ & & 0 & \\ & & & I_{r_s} \end{bmatrix}$$

$$=\lambda_1 Q_1+\lambda_2 Q_2+\cdots+\lambda_s Q_s \ =\sum_{i=1}^{s}\lambda_i Q_i \qquad P_i=PQ_iP^{-1}$$

$$\Rightarrow \ A=P\left(\sum_{i=1}^{s}\lambda_i Q_i\right)P^{-1}\ =\sum_{i=1}^{s}\lambda_i PQ_iP^{-1}\ =\sum_{i=1}^{s}\lambda_i P_i$$

133

三、可对角化矩阵的谱分解

$$Q_i=\begin{bmatrix} \ddots & & \\ & I_{r_i} & \\ & & \ddots \end{bmatrix}$$

1. 可对角矩阵的谱分解

- 结果：

$$A=\sum_{i=1}^{s}\lambda_i P_i \qquad P_i=PQ_iP^{-1}$$

幂等矩阵

Q_i，P_i具性质：

$$\sum_{i=1}^{s} Q_i=I \qquad Q_i^2=Q_i \qquad Q_iQ_j=0, i\neq j$$

$$\Rightarrow \ \sum_{i=1}^{s} P_i=I \qquad P_i^2=P_i \qquad P_iP_j=0, i\neq j$$

134

2、 矩阵可以对角化的一个充要条件

定理3.5（P.73）矩阵A可以相似对角化当且仅当矩阵A有谱分解

$$A=\sum_{i=1}^{s}\lambda_i P_i$$

满足条件：$P_i^2=P_i,\ P_iP_j=0, i\neq j,\ \sum_{i=1}^{s}P_i=I$

必要性前面已证。

充分性的证明(利用$x=Ix$及充分条件)：

在A有谱分解时 ⟶ $C^n=V_{\lambda 1}\oplus V_{\lambda 2}\oplus\ldots\oplus V_{\lambda n}$

135

3. 幂等矩阵的性质

定理3.4（P.72）$P\in F^{n\times n}$，$P^2=P$，则

- 矩阵 P^H 和矩阵（$I-P$）仍然是幂等矩阵。
- P 的谱 $\subseteq\{0, 1\}$，P 可相似于对角形。
- $F^n = N(P)\oplus R(P)$

  $N(P)=V_{\lambda=0}$，$R(P)=V_{\lambda=1}$
- P 和（$I-P$）的关系

  $N(I-P)=R(P)$，$R(I-P)=N(P)$

136

Hermite 矩阵($A^H=A$)的基本性质

Hermite 阵的特征值为实数；

Hermite 阵不同的特征值对应的特征向量正交；

对任一Hermite 阵A存在酉矩阵U使得A酉相似于对角阵（可由定理3.10得出）；

半正定（正定）Hermite 阵的特征值非负（为正）。

Hermite 矩阵的谱分解

定理3.6（P.73）设$A\in F^{n\times n}$是秩为k的半正定的Hermite 矩阵，则A可以分解为下列半正定矩阵的和：

$$A = v_1v_1^H + v_2v_2^H + \dots + v_kv_k^H,$$

其中$\{v_1, v_2, \dots, v_k\}$是F^n中的正交向量组。

137

§ 3.2 Schur 分解和正规矩阵

已知：欧氏空间中的对称矩阵A可以正交相似于对角形。

讨论：一般方阵A，在什么条件下可以酉相似于对角矩阵？

在内积空间中讨论问题，涉及：

U的列向量是空间C^n中的标准正交基

- 空间 C^n、$C^{n\times n}$，
- 酉矩阵U，$U^HU=I$，$U^{-1}=U^H$
- 酉相似：$U^HAU=J$ $\Leftrightarrow$ $U^{-1}AU=J$
- 相似关系

重点：理论结果

138

一、 Schur 分解

1、 可逆矩阵的UR分解

定理3.7（P.74）$A\in C^{n\times n}$为可逆矩阵，则存在酉矩阵U和主对角线上元素皆正的上三角矩阵R，使得A=UR。（称A=UR为矩阵A的酉分解）

证明：源于Schmidt正交化方法（P.18）

列满秩矩阵的QR分解(Th3.8($r=n$)) → Th3.7)

定理3.8（P.76）：设矩阵$A\in C^{n\times r}$是列满秩的矩阵，则矩阵A可以分解为A=QR，其中$Q\in C^{n\times r}$的列向量是标准正交的向量组，$R\in C^{r\times r}$是主对角线上元素为正数的上三角形矩阵。

例题1（例7）求矩阵A的UR分解。

139

一、 Schur 分解

列满秩矩阵的QR分解的推导

设$A=(\alpha_1, \alpha_2, \dots, \alpha_r)$，$\operatorname{rank}(A)=r$，则由Schmidt正交化方法（P.18）可将A的列化为标准正交的向量组：$(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_r)$，

记$Q=(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_r)$，则有关系式：

$$(\alpha_1, \alpha_2, \dots, \alpha_r) = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_r)\begin{bmatrix} \|\beta_1\| & (\alpha_2,\varepsilon_1) & \cdots & (\alpha_r,\varepsilon_1) \\ & \|\beta_2\| & \cdots & (\alpha_r,\varepsilon_2) \\ & & \ddots & \vdots \\ & & & \|\beta_r\| \end{bmatrix}$$

即得QR分解A=QR。

UR分解和前面某些分解结合，可导出一些新的分解。

140

2 、Schur 分解（Jordan形+UR）

定理3.9（P.76）对矩阵$A\in C^{n\times n}$，存在酉矩阵U和上三角矩阵T，使得

$$U^HAU = T = \begin{bmatrix} \lambda_1 & * & * & * \\ & \lambda_2 & \ddots & * \\ & & \ddots & * \\ & & & \lambda_n \end{bmatrix}$$

证明要点：

- $A = PJ_AP^{-1}$，
- $P = UR$，
- $A = PJ_AP^{-1} = U(RJR^{-1})U^H = UTU^H$。
- $A = UTU^H$ 称为A的Schur分解。

141

QR或UR分解的应用：其它一些分解

(1) A行满秩（可逆），则 $A=LV$，L为正线下三角矩阵，V的行标准正交（V为酉阵）；

(2) 满秩分解，$A_{m\times n}=Q_{m\times r}D_{r\times n}$，$\operatorname{rank}(A)=r$，其中Q的列标准正交，$\operatorname{rank}(D)=r$；

(3) $A_{m\times n}=U_{m\times m}\begin{pmatrix} B_r & O \\ O & O\end{pmatrix}V^H_{n\times n}$, B_r可逆, U和V为酉矩阵。

$$(1)\ A^H=UR \Rightarrow A=R^HU^H,\ or\ A^H=QR \Rightarrow A=R^HQ^H,$$

$$(2)\ A=BC \Rightarrow A=(QR)C=Q(RC)=QD,$$

$$(3)\ A=P\begin{pmatrix} I_r & O \\ O & O\end{pmatrix}Q=UR\begin{pmatrix} I_r & O \\ O & O\end{pmatrix}LV^H=U\begin{pmatrix} R_{11}L_{11} & O \\ O & O\end{pmatrix}V^H.$$

142

二、正规矩阵（Normal Matrices）

1、定义3.3 方阵A是正规矩阵 $\Leftrightarrow$ $A^HA=AA^H$。

常见的正规矩阵（P78，例9）

- 对角矩阵
- 实对称和反对称矩阵：$A^T=A$，$A^T=-A$
- Hermite矩阵和反Hermite矩阵：$A^H=A$，$A^H=-A$
- 正交矩阵和酉矩阵：$A^TA=AA^T=I$，$A^HA=AA^H=I$

例题1（*P.78，eg10*）设A为正规矩阵，B酉相似于A，证明B也是正规矩阵。

正规是酉相似的不变性质

例题2、$\forall A\in F^{m\times n}$，矩阵$A^HA$和矩阵$AA^H$是正规矩阵。

143

2、正规矩阵的基本特性

定理3.10 （P.78）

$A\in C^{n\times n}$正规 $\Leftrightarrow$ A酉相似于对角形。

- 推论：$A\in C^{n\times n}$是正规阵$\Leftrightarrow$A有n个标准正交的特征向量构成空间C^n的标准正交基。

定理3.11（P.80）（正规矩阵的谱分解）

Hermite 性

A正规 $\Leftrightarrow$ A有如下谱分解：

$$A=\sum_{i=1}^{s}\lambda_iP_i \qquad P_i^2=P_i,\ P_i^H=P_i$$

$$P_i=UQ_iU^H \qquad P_iP_j=0,\ i\neq j$$

$$\sum_{i=1}^{s}P_i=I$$

144

3、正规性质的应用举例

- 利用正规阵酉相似于对角阵，可方便地证明一些相关命题。

例题1(P.79，eg11) Hermite阵的特征值为实数;

例题2(P.79，eg12) 酉矩阵的特征值的模为1;

例题3 设$A\in R^{n\times n}$，$A^T=-A$，证明

1. A的特征值是零和纯虚数。
2. 矩阵A的秩是偶数。

实系数多项式的复数根成对出现。

$$g(\lambda)=\sum_{k=0}^{m}a_k\lambda^k=0 \Leftrightarrow g(\bar{\lambda})=\sum_{k=0}^{m}a_k\bar{\lambda}^k=0$$

145

§ 3.3 矩阵的奇异值分解

Singular value decomposition

(SVD)

146

§ 3.3 矩阵的奇异值分解

- 概述：
  - 矩阵的奇异值分解是酉等价型的分解: $A\in C^{m\times n}$，存在酉矩阵$U\in C^{m\times m}$, $V\in C^{n\times n}$,使得 $A=U\Sigma V^H$。
  - 矩阵A等价于$\Sigma=\begin{bmatrix}\Delta_r & 0\\ 0 & 0\end{bmatrix}_{m\times n}$ $\quad \Delta_r=\begin{bmatrix}\sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r\end{bmatrix}$
  - 奇异值分解基本适用于<u>内积空间</u>中与矩阵秩相关的问题
  - A的奇异值分解依赖于正规矩阵A^HA的酉相似分解。

147

一、矩阵A的奇异值及其性质

1、矩阵A^HA和AA^H的性质：$A\in C^{m\times n}$

$A^HA\in C_{n\times n}$, $AA^H\in C_{m\times m}$ 为Hermite矩阵。

定理3.12（P.82）

1. 秩(A) = 秩(A^HA) = 秩(AA^H)。（利用酉等价）
2. A^HA 和 AA^H 的非零特征值相等。
3. A^HA 和 AA^H 半正定。
   r(A) = n时, A^HA 正定; r(A) = m时, AA^H 正定。

⇒ A^HA和AA^H 的特征值是非负实数：$\lambda_1\geq \lambda_2\geq\ldots\geq\lambda_n\geq 0$

2、奇异值的定义（P.72）

$A\in C^{m\times n}$，秩(A) = r，设A^HA的特征值$\lambda_1\geq\lambda_2\geq\ldots\geq\lambda_r>0$，$\lambda_{r+1}=\lambda_{r+2}=\ldots=\lambda_n=0$，则矩阵A的奇异值

$$\sigma_i=\sqrt{\lambda_i},\quad i=1,2,\ldots,r.$$

148

3、特殊矩阵的奇异值

- 定理3.13（P.82）：
  - 正规矩阵A的奇异值等于A的(非零)特征值的模。
  - 正定的Hermite矩阵A的奇异值就是A的特征值。
  - 酉等价矩阵的奇异值相等。

> A和B酉等价，则A^HA和B^HB酉相似。
> 奇异值是酉等价的不变性质。

$$A=U\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}U^H \Rightarrow A^H=U\begin{bmatrix}\bar{\lambda}_1 & & & \\ & \bar{\lambda}_2 & & \\ & & \ddots & \\ & & & \bar{\lambda}_n\end{bmatrix}U^H \Rightarrow A^HA=U\begin{bmatrix}|\lambda_1|^2 & & & \\ & |\lambda_2|^2 & & \\ & & \ddots & \\ & & & |\lambda_n|^2\end{bmatrix}U^H$$

$$\Rightarrow \sigma_i=\sqrt{|\lambda_i|^2}=|\lambda_i|, \lambda_i\neq 0, 1\leq i\leq n.$$

149

3、特殊矩阵的奇异值

- 定理3.13（P.82）：
  - 正规矩阵A的奇异值等于A的(非零)特征值的模。
  - 正定的Hermite矩阵A的奇异值就是A的特征值。
  - 酉等价矩阵的奇异值相等。

> A和B酉等价，则A^HA和B^HB酉相似。
> 奇异值是酉等价的不变性质。

$$A=U\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}U^H \Rightarrow A^H=U\begin{bmatrix}\bar{\lambda}_1 & & & \\ & \bar{\lambda}_2 & & \\ & & \ddots & \\ & & & \bar{\lambda}_n\end{bmatrix}U^H \Rightarrow A^HA=U\begin{bmatrix}|\lambda_1|^2 & & & \\ & |\lambda_2|^2 & & \\ & & \ddots & \\ & & & |\lambda_n|^2\end{bmatrix}U^H$$

$$\Rightarrow \sigma_i=\sqrt{|\lambda_i|^2}=|\lambda_i|=\lambda_i, \because \lambda_i>0, 1\leq i\leq n.$$

150

3、特殊矩阵的奇异值

- 定理3.13（P.82）：
  - 正规矩阵A的奇异值等于A的(非零)特征值的模。
  - 正定的Hermite矩阵A的奇异值就是A的特征值。
  - 酉等价矩阵的奇异值相等。

> A和B酉等价，则A^HA和B^HB酉相似。
> 奇异值是酉等价的不变性质。

$$A=UBV^H \quad \Rightarrow A^H=VB^HU^H \quad \Rightarrow A^HA=VB^HBV^H$$

$$\Rightarrow AA^H=UBB^HU^H$$

151

二、矩阵的奇异值分解

1、定理3.14（P.83）

- 任何矩阵$A\in C^{m\times n}$，秩(A) = r，则存在酉矩阵$U\in C^{m\times m}$，$V\in C^{n\times n}$，使得

$$A=U\left[\begin{array}{cccc|c}\sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & 0\\ & & & \sigma_r & \\ \hline & & 0 & & 0\end{array}\right]V^H \qquad \Delta=\begin{bmatrix}\sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r\end{bmatrix}$$

- 证明思想：$AV=U\Sigma$，即 $Av_i=\sigma_iu_i$，$i<=r$；$=0$, $i>r$。
  - A^HA正规，$V^HA^HAV=\begin{bmatrix}\Delta_r^2 & \\ & 0\end{bmatrix}$，⇒ 酉矩阵V。
- 令 $u_i=\dfrac{Av_i}{\sigma_i}$ ，i=1，2，…，r，得$U_1=[u_1, u_2, \ldots, u_r]$
  将其扩充为标准正交基⇒ 酉矩阵U。

152

- 例题1 求矩阵A的奇异值分解，$A=\begin{bmatrix}1 & 1\\ 0 & 1\\ 1 & 0\end{bmatrix}$。

解 (1) 求

$$A^HA=\begin{bmatrix}1 & 0 & 1\\ 1 & 1 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 0 & 1\\ 1 & 0\end{bmatrix}=\begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix}$$

的特征值：

$$|\lambda I-A^HA|=\begin{bmatrix}\lambda-2 & -1\\ -1 & \lambda-2\end{bmatrix}=(\lambda-3)(\lambda-1)=0$$

$$\Rightarrow \lambda_1=3,\lambda_2=1 \quad \Rightarrow \sigma_1=\sqrt{3},\sigma_2=1.$$

(2) 求 λ_i 对应的特征向量，确定V：

$$\lambda_1=3: x_1-x_2=0 \Rightarrow X_1=k_1(1,1)^T;$$

$$\lambda_2=1: x_1+x_2=0 \Rightarrow X_2=k_2(-1,1)^T.$$

取 $\alpha_1=(1,1)^T,\alpha_2=(-1,1)^T$.

153

例题1 求矩阵A的奇异值分解，A=$\begin{bmatrix}1&1\\0&1\\1&0\end{bmatrix}$。

解 (1) 求 $A^HA=\begin{bmatrix}2&1\\1&2\end{bmatrix}$ 的特征值：

$$\lambda_1=3,\lambda_2=1\Rightarrow\sigma_1=\sqrt{3},\sigma_2=1.$$

(2) 求 λ_i 对应的特征向量，确定V：

取 $\alpha_1=(1,1)^T,\alpha_2=(-1,1)^T$，已正交！

标准化得V：$v_1=\frac{1}{\sqrt{2}}(1,1)^T,v_2=\frac{1}{\sqrt{2}}(-1,1)^T,$

(3) 求U：$u_1=Av_1/\sigma_1,u_2=Av_2/\sigma_2$

$\Rightarrow u_1=(2,1,1)^T/\sqrt{6},u_2=(0,1,-1)^T/\sqrt{2}$

u_3？ 扩充 $u_3\perp u_1,u_2$ $\Rightarrow A=U\Sigma V^H$ Σ？

154

例题2（P.84，eg13）

求矩阵A和B的奇异值分解，A=$\begin{bmatrix}1&0&1\\0&1&1\\0&0&0\end{bmatrix}$；

B = $\begin{bmatrix}1&1&0\\0&0&1\end{bmatrix}$。

解 (1) 求A^HA的特征值：$\lambda_1=3,\lambda_2=1,\lambda_3=0$

$$\Rightarrow\sigma_1=\sqrt{3},\sigma_2=1.$$

求λ_i的特征向量(已正交), 正交化定出V：

$\alpha_1=(1,1,2)^T,\alpha_2=(1,-1,0)^T,\alpha_3=(1,1,-1)^T,$

$\Rightarrow v_1=(1,1,2)^T/\sqrt{6},v_2=(1,-1,0)^T/\sqrt{2},v_3=(1,1,-1)^T/\sqrt{3}.$

$\Rightarrow u_1=Av_1/\sigma_1,u_2=Av_2/\sigma_2$

155

例题2（P.84，eg13）

求矩阵A和B的奇异值分解，A=$\begin{bmatrix}1&0&1\\0&1&1\\0&0&0\end{bmatrix}$；

B = $\begin{bmatrix}1&1&0\\0&0&1\end{bmatrix}$。

解 (1) 求A^HA的特征值：$\lambda_1=3,\lambda_2=1,\lambda_3=0$

$$\Rightarrow\sigma_1=\sqrt{3},\sigma_2=1.$$

求λ_i的特征向量(已正交), 正交化定出V：

$v_1=(1,1,2)^T/\sqrt{6},v_2=(1,-1,0)^T/\sqrt{2},v_3=(1,1,-1)^T/\sqrt{3}.$

$\Rightarrow u_1=(1,1,0)^T/\sqrt{2},u_2=(1,-1,0)^T/\sqrt{2},$ u_3？

扩充 $u_3\perp u_1,u_2\Rightarrow u_3=(0,0,1)^T$ $\Rightarrow A=U\Sigma V^H$

156

例题2（P.84，eg13）

求矩阵A和B的奇异值分解，A=$\begin{bmatrix}1&0&1\\0&1&1\\0&0&0\end{bmatrix}$；

B = $\begin{bmatrix}1&1&0\\0&0&1\end{bmatrix}$。

解 (2) 求B^HB的特征值：$\lambda_1=2,\lambda_2=1,\lambda_3=0$

$$\Rightarrow\sigma_1=\sqrt{2},\sigma_2=1.$$

求λ_i的特征向量(已正交), 正交化定出V：

$v_1=(1,1,0)^T/\sqrt{2},v_2=(0,0,1)^T,v_3=(-1,1,0)^T/\sqrt{2}.$

$\Rightarrow u_1=Bv_1/\sigma_1=(1,0)^T,u_2=Bv_2/\sigma_2=(0,1)^T$，已得U！

$$\Rightarrow B=U\Sigma V^H$$

157

2、矩阵U，V的空间性质：

右奇异向量

- V=[v_1，v_2，…，v_r，…，v_n] =[V_1 V_2]∈$C^{n\times n}$的列向量是空间C^n的标准正交基。
  - V_2的列向量是空间N(A)的标准正交基（AV_2=0）。
- U=[u_1，u_2，…，u_r，…，u_m] =[U_1 U_2]∈$C^{m\times m}$的列向量是空间C^m的标准正交基。
  - U_1 的列向量是R(A)的标准正交基（$A=U_1\Delta_rV_1^H$）。

左奇异向量

3、奇异值分解的展开形式及其应用 $A=U_1\Delta_rV_1^H$

- 定理3.15（P.87）（由奇异值分解展开得到！）

$$A=\sigma_1u_1v_1^H+\sigma_2u_2v_2^H+\cdots+\sigma_ru_rv_r^H$$

158

例题：图像的数字化技术与矩阵的奇异值分解

- 计算机处理图像技术的第一步是图像的数字化存储技术，即将图像转换成矩阵来存储。
- 转换的原理是将图形分解成象素（pixels）的一个矩形的数阵，其中的信息就可以用一个矩阵$A=(a_{ij})_{m\times n}$来存储。矩阵A的元素a_{ij}是一个正的数，它相应于象素的灰度水平（gray level）的度量值。
- 由于一般来讲，相邻的象素会产生相近的灰度水平值，因此有可能在满足图像清晰度要求的条件下，将存储一个m×n阶矩阵需要存储的m×n个数减少到n+m+1的一个倍数。

159

- 压缩数字化图形存储量的方法主要是应用矩阵的奇异值分解和矩阵范数下的逼近。如果图象的数字矩阵A的奇异值分解为：A=UΣV^T，其展开式：

$$A=\sigma_1u_1v_1^H+\sigma_2u_2v_2^H+\cdots+\sigma_ru_rv_r^H$$

- 压缩矩阵A的方法是取一个秩为k（$k\le r$）的矩阵A_k来逼近矩阵A。
- A_k按如下方法选取：

$$A_k=\sigma_1u_1v_1^H+\sigma_2u_2v_2^H+\cdots+\sigma_ku_kv_k^H$$

- 在秩为k（$k\le$n）的所有矩阵中，矩阵A_k所对应的图象和矩阵A所对应的图象最相近。一般的，k越大图象就越清晰。经典的方法是选取接近k，使A_k的存储量比A的存储量减少20%。

160

- 存储矩阵A_k只需要存储k个奇异值，k个m维向量u_i和n维向量v_j的所有分量，共计k（m+n+1）个元素。
- 如果m=n=1000，存储原矩阵A需要存储1000×1000个元素。取k=100时，图象已经非常清晰了，这时的存储量是100（2000+1）=200100个数。
- 和矩阵A比较，存储量减少了80%。

161

图像数据的奇异值分解压缩：秩从4到128

162

三、矩阵的奇异值分解和线性变换T_A

- 矩阵$A\in C^{m\times n}$可以定义线性变换

$$T_A:\quad C^n \to C^m$$

- 设矩阵的奇异值分解$A=U\Sigma V^H$，则将U和V的列分别取做空间C^m、C^n的基，则变换T_A的矩阵为Σ: $AV = U\Sigma$，进而有，

$$\forall \alpha=VX \in C^n, \text{则} T_A\alpha=(U\Sigma V^H)VX=U(\Sigma X)= U\begin{bmatrix}\sigma_1 x_1\\ \sigma_2 x_2\\ \vdots\\ \sigma_r x_r\\ 0\end{bmatrix}$$

- 对实矩阵$A_{m\times n}$，变换T_A在单位球上的象是R^n中球面(r=n)或椭球体(r<n)：定理3.16（P.88）
- 例14（来自例13的奇异值分解）

163

四、矩阵的极分解（Polar Decomposition）

- 方阵的极分解
  - 设矩阵$A\in C^{n\times n}$，则矩阵A的奇异值分解:
    $A = U\Sigma V^H = U\Sigma(U^HU)V^H = (U\Sigma U^H)UV^H = PQ$
  - P相似于Σ，是半正定的Hermite 矩阵。
  - Q是酉矩阵
- 定理3.17 （P.89） $A = PQ$
- 方阵极分解的意义和应用 ~复数表示$z = re^{i\theta}$
  - 描述变换Y=AX的旋转(Q)和拉伸(P)

164

- 例题1 （ P.90） 求矩阵$A=\begin{bmatrix}\sqrt{3} & 2\\ 0 & \sqrt{3}\end{bmatrix}$的极分解，依此讨论变换Y=AX的几何特性。

解

$$A = PQ = \begin{bmatrix}\frac{5}{2} & \frac{\sqrt{3}}{2}\\ \frac{\sqrt{3}}{2} & \frac{5}{2}\end{bmatrix}\begin{bmatrix}\frac{\sqrt{3}}{2} & \frac{1}{2}\\ -\frac{1}{2} & \frac{\sqrt{3}}{2}\end{bmatrix}$$

165

第4章 矩阵的广义逆

The Pseudoinverse

166

矩阵的广义逆

- 概述：
- 矩阵的逆：$A_{n\times n}$，$\exists B_{n\times n}$，$BA= AB = I$，则 $B=A^{-1}$
- 广义逆的目标：推广逆的概念
  - 对一般的矩阵 $A_{m\times n}$可建立逆的部分性质。
  - 当矩阵$A_{n\times n}$可逆时，广义逆与逆相一致。
  - 应用：广义逆可以作为方程组AX=b求解和最小二乘法的理论分析工具。

若A可逆，推出：$BA=I$，$AB = I$，进而有

$ABA = A$，$BAB = B$，$(AB)^H = AB$，$(BA)^H = BA$，

由此可引出多种广义逆。这里重点讨论三种：单侧逆，减号逆和加号逆。

167

§ 4. 1 矩阵的左逆与右逆

一、满秩矩阵和单侧逆

必要条件

右逆存在 $\Rightarrow m=r(AC)\le r(A)\le n$

左逆存在 $\Rightarrow n=r(BA)\le r(A)\le m$

1、左逆和右逆的定义

- 定义4.1 (*P.93*)
  - $A \in C^{m\times n}$，$\exists B \in C^{n\times m}$，$BA=I_n$，则称矩阵$B$为矩阵$A$的左逆，记为 $B = A_L^{-1}$
  - $A \in C^{m\times n}$，$\exists C \in C^{n\times m}$，$AC=I_m$，则称矩阵$C$为矩阵A的右逆，记为 $C = A_R^{-1}$

例题1 求矩阵A的左逆：$A = \begin{bmatrix}-1 & 0\\ 0 & 1\\ 2 & -1\end{bmatrix}$ $A_L^{-1} = \begin{bmatrix}-1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}$

A的右逆？不存在！

168

§ 4. 1 矩阵的左逆与右逆

一、满秩矩阵和单侧逆

必要条件

右逆存在 $\Rightarrow m=r(AC)\le r(A)\le n$

左逆存在 $\Rightarrow n=r(BA)\le r(A)\le m$

1、左逆和右逆的定义

- 定义4.1 (*P.93*)
  - $A \in C^{m\times n}$，$\exists B \in C^{n\times m}$，$BA=I_n$，则称矩阵$B$为矩阵$A$的左逆，记为 $B = A_L^{-1}$
  - $A \in C^{m\times n}$，$\exists C \in C^{n\times m}$，$AC=I_m$，则称矩阵$C$为矩阵A的右逆，记为 $C = A_R^{-1}$

例题1 求矩阵A的左逆：$A = \begin{bmatrix}-1 & 0\\ 0 & 1\\ 0 & 0\end{bmatrix}$ $A_L^{-1} = \begin{bmatrix}-1 & 0 & a\\ 0 & 1 & b\end{bmatrix}$

左逆不唯一！

169

2、左逆和右逆存在的条件

$n=r(BA)\le r(A)\le m,\ n$

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

- A_L^{-1}的存在性
  - 直观分析
  - A_L^{-1} 存在 ⇔ 矩阵A列满秩
  - $A_L^{-1} = (A^HA)^{-1}A^H$
- 定理4.1(*P.93*) 设$A \in C^{m\times n}$，下列条件等价

1. A左可逆； $BA = I_n$
2. A的零空间 $N(A) = \{0\}$； $Ax = 0 \Longrightarrow x = BAx = 0$
3. $m \ge n$，秩$(A) = n$，即A是列满秩的； $n\text{-}r(A) = 0$
4. 矩阵A^HA可逆，且 $A_L^{-1} = (A^HA)^{-1}A^H$。 $r(A^HA) = r(A)$

如前例 矩阵 $A = \begin{bmatrix}-1 & 0\\ 0 & 1\\ 2 & -1\end{bmatrix}$ 左可逆，A^T右可逆。

如何求左或右逆？ 可用行或列初等变换！

170

$m=r(AC)\le r(A)\le m,\ n$

$AA^H(AA^H)^{-1}=I$

- 矩阵右逆的存在性
  - 定理4.2 (*P.94*) 设$A \in C^{m\times n}$，则下列条件等价：

1. 矩阵A右可逆； $AC = I_m$
2. A的列空间 $R(A) = C^m$； $x = ACx \Longrightarrow x \in R(A)$
3. $n \ge m$, 秩$(A) = m$，即A是行满秩的； $r(A)=dimR(A)$
4. 矩阵 AA^H 可逆，且 $r(AA^H) = r(A)$

$$A_R^{-1} = A^H(AA^H)^{-1}$$

讨论：可逆矩阵$A_{n\times n}$的左、右逆和逆的关系

- 可逆矩阵A的左、右逆就是矩阵A的逆A
- $A^{-1} = (A^HA)^{-1}A^H = A^H(AA^H)^{-1}$

171

二、单侧逆和求解线性方程组AX=b

- 讨论
  - 方程组AX=b 有解与左、右逆存在的关系。
  - 借助于左、右逆求AX=b的形如X=Bb的解。
- 1、右可逆矩阵
  - 定理4.4 (*P.95*)
  1. $A\in C^{m\times n}$右可逆，则$\forall b\in C^m$，AX=b 有解。
  2. $X=A_R^{-1}b$ 是方程组 AX=b 的解，特别地，$X=A^H(AA^H)^{-1}b$ 是一个解。

  由AC=I，知ACb=Ib=b，又AA^H可逆，得证。

172

二、单侧逆和求解线性方程组AX=b

- 2、左可逆矩阵
  - 求解分析：
  - 定理4.3 (*P.94*) 设矩阵$A\in C^{m\times n}$左可逆，B是矩阵A的任何一个左逆，则
  1. AX=b有形如X=Bb的解的充要条件是 $(I_m-AB)b=0$ (¤)
  2. 当(¤)式成立时，方程组的解是惟一的，而且惟一解是$X=(A^HA)^{-1}A^Hb$

  证明：1. b=AX=ABAX=ABb；2. $X=(A^HA)^{-1}A^HAX$

讨论：对任何满足式(¤)的左逆B，X=Bb都是方程组的解，如何解释方程组的解是惟一的？

173

§ 4.2 广义逆矩阵

- 思想：用公理来定义广义逆。
- 一、减号广义逆
  - 定义4.2 (*P.95*) $A\in C^{m\times n}$，如果，$\exists G\in C^{n\times m}$使得AGA=A，则称矩阵G为A的减号广义逆，或{1}-逆。
  - A的减号逆集合记为$A\{1\}=\{A_1^-, A_2^-, \ldots, A_k^-\}$

    例题1 $A\in C^{n\times n}$可逆，则$A^{-1}\in A\{1\}$；

    A单侧可逆，则$A^{-1}_L\in A\{1\}$；$A^{-1}_R\in A\{1\}$。
  - 若$A=0\in C^{m\times n}$，则$A\{1\}=C^{n\times m}$。

174

§ 4.2 广义逆矩阵

- 减号逆的求法：初等变换求等价标准型

定理4.5(*P.95*)设$A\in C^{m\times n}$，rank(A) = r，若存在可逆阵P，Q 使 $PAQ=\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}$，则 $G\in A\{1\}$ 当且仅当

$$G=Q\begin{pmatrix} I_r & U\\ V & W\end{pmatrix}P,$$ 其中U，V，W任意。

- 证明思路：令 $G=Q(Q^{-1}GP^{-1})P=Q\begin{pmatrix} X & U\\ V & W\end{pmatrix}P$

  由 AGA = A 可推出：$X=I_r$。

任一矩阵的减号逆总存在，且一般不惟一！

175

§ 4.2 广义逆矩阵　　$AA^-A=A$

- 减号逆的性质：定理4.6 - 定理4.8

定理4.6 (*P.96*)设$A\in C^{m\times n}$，则A的{1}-逆惟一当且仅当m = n，且A^{-1}存在（即A可逆）。

定理4.7 (*P.96*)设$A\in C^{m\times n}$，则A^-满足

(1) rank(A) <= rank(A^-);

(2) AA^-与A^-A都是幂等阵，且

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

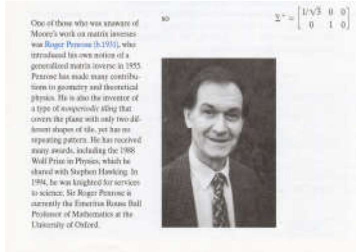
(3) $R(AA^-)=R(A)$, $N(A^-A)=N(A)$。

定理4.8 (*P.97*)设$A\in C^{m\times n}$, $A^-\in A\{1\}$。若AX = b有解，则其通解可表示为：$X=A^-b+(I_n-A^-A)Z$，$Z\in C^n$任意。

A^-b为AX=b的特解, $(I_n-A^-A)Z$为AX=0的通解.

176

E.H. Moore and Roger Penrose



177

二、Moore-Penrose (M-P) 广义逆

- 由Moore 1920年提出，1955年由Penrose独立研究和发展。

1、定义4.3 (*P.98*) 设矩阵 $A\in C^{m\times n}$，如果 $\exists G\in C^{n\times m}$，使得
  1. AGA = A
  2. GAG = G
  3. $(AG)^H=AG$
  4. $(GA)^H=GA$

$A=G^+=(A^+)^+$

$(A^+)^H=(A^H)^+$

则称G为A的M-P广义逆，记为$G=A^+$ (简称加号逆)。

例题2 讨论原有的逆的概念和M-P广义逆的关系。

$A^{-1}=A^+$；$A^{-1}_L=(A^HA)^{-1}A^H=A^+$；

$A^{-1}_R=A^H(AA^H)^{-1}=A^+$；若$\exists A^+$，则$A^+$是A{1}。

178

- 2、M-P广义逆的惟一性
  - 定理4.9 (P.98) 如果A有M-P广义逆，则A的M-P广义逆是惟一的。
- 3、M-P广义逆的存在性及其求法
  - 定理4.10（P.99）任何矩阵都有M-P广义逆。
  - 求法：
    - 设A有满秩分解 A = BC，则有 $A^+=C^H(CC^H)^{-1}(B^HB)^{-1}B^H$。
    - (定理4.11) 设A有奇异值分解：

$$A=U\begin{bmatrix}\Delta_r & 0\\ 0 & 0\end{bmatrix}V^H,\ \text{则}\quad A^+=V\begin{bmatrix}\Delta_r^{-1} & 0\\ 0 & 0\end{bmatrix}U^H$$

179

- 4、M-P广义逆的性质
- 定理4.12 (P.100)：A^+满足下列性质：
  1. $(A^+)^+=A$
  2. $(A^+)^H=(A^H)^+$
  3. $(\lambda A)^+=\lambda^+A^+$
  4. A 列满秩，则 $A^+=(A^HA)^{-1}A^H$，左逆

     A 行满秩，则 $A^+=A^H(AA^H)^{-1}$；右逆
  5. A 有满秩分解：A = BC，则 $A^+=C^+B^+$。

  $\mathrm{rank}(A)=\mathrm{rank}(A^+)$; $\mathrm{rank}(A)=\mathrm{rank}(A^+A)=\mathrm{rank}(AA^+)$.

A^+与A^{-1}性质的差异比较：

$(AB)^{-1}=B^{-1}A^{-1}$，一般不成立 $(AB)^+=B^+A^+$（只有满秩分解成立）

$(A^{-1})^k=(A^k)^{-1}$，但不成立$(A^+)^k=(A^k)^+$

180

例题1 求下列特殊矩阵的广义逆；

- 零矩阵0； $0^+_{m\times n} = 0_{n\times m}$
- 1阶矩阵(数) a； $a^+ = \begin{cases}(a^Ha)^{-1}a^H = a^{-1}, & a \neq 0;\\ 0, & a = 0.\end{cases}$
- 对角矩阵Λ

$$\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}^+ = \begin{bmatrix}\lambda_1^+ & & & \\ & \lambda_2^+ & & \\ & & \ddots & \\ & & & \lambda_n^+\end{bmatrix}$$

例题2 求非零向量 $x = \begin{bmatrix}x_1\\x_2\\ \vdots \\ x_n\end{bmatrix}$ 的M-P广义逆。

$$x^+ = \frac{x^H}{x^Hx} = \frac{x^H}{\|x\|^2}$$

单位向量： $x^+ = x^H$

181

例题3 $x \in F^n, y \in F^m, x \neq 0, y \neq 0,$ $\quad x^+ = \frac{x^H}{x^Hx} = \frac{x^H}{\|x\|^2}$

$$(xy^H)^+ = (y^H)^+x^+ = \frac{y}{\|y\|^2}\frac{x^H}{\|x\|^2} = \frac{(xy^H)^H}{\|x\|^2\|y\|^2}$$

$$\begin{pmatrix}1\\0\end{pmatrix}^+ = (1,0);\ (1,0)^+ = \begin{pmatrix}1\\0\end{pmatrix};\ \begin{pmatrix}1\\1\end{pmatrix}^+ = \frac{1}{2}(1,1);\ (1,1)^+ = \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}$$

$$\begin{pmatrix}1&1\\0&0\end{pmatrix}^+ = \left(\begin{pmatrix}1\\0\end{pmatrix}(1,1)\right)^+ = (1,1)^+\begin{pmatrix}1\\0\end{pmatrix}^+ = \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}(1,0) = \frac{1}{2}\begin{pmatrix}1&0\\1&0\end{pmatrix}$$

$$\begin{pmatrix}1&0\\0&0\end{pmatrix}^+ = \left(\begin{pmatrix}1\\0\end{pmatrix}(1,0)\right)^+ = (1,0)^+\begin{pmatrix}1\\0\end{pmatrix}^+ = \begin{pmatrix}1\\0\end{pmatrix}(1,0) = \begin{pmatrix}1&0\\0&0\end{pmatrix}$$

$$\begin{pmatrix}0&1&0\\0&0&0\end{pmatrix}^+ = \left(\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}0&1&0\end{pmatrix}\right)^+ = (0,1,0)^+\begin{pmatrix}1\\0\end{pmatrix}^+ = \begin{pmatrix}0\\1\\0\end{pmatrix}(1,0) = \begin{pmatrix}0&0\\1&0\\0&0\end{pmatrix}$$

182

例题4 $\begin{pmatrix}I_r\\0\end{pmatrix}^+ = (I_r,0);\ (I_r,0)^+ = \begin{pmatrix}I_r\\0\end{pmatrix};\ A_r(\in F^{r\times n})$行满秩，$\begin{pmatrix}A_r\\0\end{pmatrix}^+ = ?$

$$\begin{pmatrix}A_r\\0\end{pmatrix}^+ = \left(\begin{pmatrix}I_r\\0\end{pmatrix}A_r\right)^+ = A_r^+(I_r,0) = (A_r^+,0)$$

$$\begin{pmatrix}1&1\\0&0\end{pmatrix}^+ = \begin{pmatrix}(1,1)^+ & \begin{matrix}0\\0\end{matrix}\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1&0\\1&0\end{pmatrix};\quad \begin{pmatrix}1&0\\0&0\end{pmatrix}^+ = \begin{pmatrix}(1,0)^+ & \begin{matrix}0\\0\end{matrix}\end{pmatrix} = \begin{pmatrix}1&0\\0&0\end{pmatrix}$$

$$\begin{pmatrix}0&1&0\\0&0&0\end{pmatrix}^+ = \begin{pmatrix}(0\ 1\ 0)^+ & \begin{matrix}0\\0\\0\end{matrix}\end{pmatrix} = \begin{pmatrix}0&0\\1&0\\0&0\end{pmatrix}$$

例题5 设 $A = \begin{bmatrix}1&1&2\\0&2&2\\1&0&1\end{bmatrix}$，求A⁺。 $A = \begin{bmatrix}1&1\\0&2\\1&0\end{bmatrix}\begin{bmatrix}1&0&1\\0&1&1\end{bmatrix}$

183

§ 4. 3 投影变换（为讨论A^+的应用做准备）

一、投影变换和投影矩阵

- 定义4.4 (*P.101*) 设 $C^n = L\oplus M$，向量 $x \in C^n$, $x = y + z$，$y\in L$，$z \in M$，如果线性变换 σ：$C^n\to C^n$，$\sigma(x) = y$， 则称σ为从 C^n沿子空间M到子空间L的投影变换。

> R(σ) = L；N(σ) = M， ⇒ $C^n = R(\sigma) \oplus N(\sigma)$
>
> L和M是σ的不变子空间；$\sigma|_L = I$；$\sigma|_M = 0$

- 投影变换的矩阵 $\sigma \xleftrightarrow{\exists\{\alpha_1,\alpha_2,\cdots,\alpha_n\}} \begin{bmatrix}I_r & 0\\ 0 & 0\end{bmatrix}$
- 投影的矩阵和变换性质：
  1. 定理4.13 (P.101) σ是投影变换 ⇔ σ是幂等变换
  2. 推论：σ为投影变换的充要条件是变换矩阵是幂等矩阵

184

二、正交投影和正交投影矩阵

1. 正交投影的定义：

定义4.5（*P.103*）设σ：$C^n \to C^n$ 是投影变换，
$C^n = R(\sigma) \oplus N(\sigma)$，如果
$R^\perp(\sigma) = N(\sigma)$，则称σ为正交投影变换。

2 正交投影矩阵 定理4.14（*P.103*）σ是正交投影 ⇔ 投影矩阵A满足：$A^2 = A, A^H = A$。

充分性: 证$R^\perp(A) = N(A)$; 必要性: 证$R(A) = R(A^H)$, $N(A)=N(A^H)$。

例题1 设W是 C^n 的子空间，证明：存在到W的投影变换，使 R(σ) = W。
类似地：在内积空间 C^n 中，存在到W的正交投影变换，使 R(σ) = W。

185

投影矩阵和正交投影矩阵的求法

投影矩阵的求法 设A：$C^n \to L$ 是投影阵，
$C^n = L \oplus M$，$\dim(L) = r$。取L和M的基
$\{y_1, y_2, \dots, y_r\}$, $\{z_1, z_2, \dots, z_{n-r}\}$，则有

$$A(y_1, y_2, \dots, y_r) = (y_1, y_2, \dots, y_r),$$
$$A(z_1, z_2, \dots, z_{n-r}) = (0, 0, \dots, 0)。$$

记 $B = (y_1, y_2, \dots, y_r)$, $C = (z_1, z_2, \dots, z_{n-r})$，则
$A(B, C) = (B, 0)$，推出 $A = (B, 0)(B, C)^{-1}$。
C^n到M的投影阵$\tilde{A}$ = ? $\qquad \tilde{A} = I - A$

例（P102例6）$R^2 = L\{(1,0)^T\} \oplus L\{(1,-1)^T\}$，$R^2$到 $L\{(1,0)^T\}$和$L\{(1,-1)^T\}$的投影阵分别为：

$$A = \begin{bmatrix}1&0\\0&0\end{bmatrix}\begin{bmatrix}1&1\\0&-1\end{bmatrix}^{-1} = \begin{bmatrix}1&1\\0&0\end{bmatrix} \qquad \tilde{A} = I_2 - A = \begin{bmatrix}0&-1\\0&1\end{bmatrix}$$

186

投影矩阵和正交投影矩阵的求法

正交投影矩阵的求法 在上述推导中，令$M = L^\perp$，
$B^HC = 0$，则

$$A = (B, 0)(B, C)^{-1}$$
$$= (B, 0)[(B, C)^H(B, C)]^{-1}(B, C)^H = B(B^HB)^{-1}B^H$$

同理，C^n到$L^\perp$的正交投影阵

$$\tilde{A} = I - A = I - B(B^HB)^{-1}B^H$$
$$= C(C^HC)^{-1}C^H$$

B 或 C 的列标准正交时，如何？ $B^HB=I_r$, $C^HC=I_{n-r}$

例(P105例7) R^3，$L = L\{a_1,a_2\}$，求到L的正交投影阵A及Ax。$\tilde{A}$?

例(P24例30) R^n上(沿u)的正交投影P变换：$P(x) = x - (x, u)u$，u 是单位向量。 $P(x) = (I_n - uu^T)x = (I_n - u(u^Tu)^{-1}u^T)x$。

187

3. 正交投影的性质

定理4.16（P.104）设W是 C^n 的子空间，$x_0\in C^n$，$x_0 \notin W$，如果σ是空间 C^n 向空间W的正交投影，则 $\|\sigma(x_0) - x_0\| \le \|y - x_0\|,\ \forall y \in W$

> 含义：点 $\sigma(x_0)$ 是空间 W 中与点 x_0距离最近的点。

证 由 $C^n = W \oplus W^\perp = R(\sigma) \oplus N(\sigma)$，知对 $y\in W$, 有 $y - \sigma(x_0) \in W$, $\sigma(x_0) - x_0 \in W^\perp$, 因此，

$$\|y - x_0\|^2 = \|y - \sigma(x_0) + \sigma(x_0) - x_0\|^2$$
$$= \|y - \sigma(x_0)\|^2 + \|\sigma(x_0) - x_0\|^2$$
$$\ge \|y - \sigma(x_0)\|^2,\ \forall y \in W$$

188

4. A^+A与AA^+的性质

定理4.15（*P.104*）Th4.14 + Th4.7

- A^+A 的性质：
  - $(A^+A)^2 = A^+A$，$(A^+A)^H = A^+A$
  - $C^n = R(A^+) \oplus N(A)$ $\qquad C^n = R(A^+A) \oplus N(A^+A)$
  - $R^\perp(A^+) = N(A)$ $\qquad R(A^+A)=R(A^+)$, $N(A^+A)=N(A)$
- AA^+ 的性质：
  - $(AA^+)^2 = AA^+$，$(AA^+)^H = AA^+$
  - $C^m = R(A) \oplus N(A^+)$ $\qquad C^m = R(AA^+) \oplus N(AA^+)$
  - $R^\perp(A) = N(A^+)$ $\qquad R(AA^+)=R(A)$, $N(AA^+)=N(A^+)$

含义：

> A^+A 是正交投影，它将向量 x 投影到空间$R(A^+)$中。
> AA^+ 是正交投影，它将向量 x 投影到空间R(A)中。

189

4.4 最佳最小二乘解

- 一、最佳最小二乘解
- $A_{m\times n}x_{n\times 1}=b_{m\times 1}$ ⟺ 有解 $\Leftrightarrow b\in R(A)$；无解 $\Leftrightarrow b\notin R(A)$

1、AX = b 的最佳最小二乘解

定义4.6（*P. 105*）

u 是最小二乘解 $\Leftrightarrow \|Au-b\|\le\|Ax-b\|,\ \forall x\in C^n$

x_0 是最佳最小二乘解 $\Leftrightarrow \|x_0\|_2\le\|u\|_2$

2、 Ax = b的最佳最小二乘解的计算

定理4.17 $x_0=A^+b$ 是 $Ax=b$ 的最佳最小二乘解。

证明思路：利用AA⁺：x_0是最小二乘解；对任一最小二乘解u有：$u-x_0\in N(A)$，从而 $x_0\perp(u-x_0)$，因此 $\|u\|_2\ge\|x_0\|_2$

190

4.4 最佳最小二乘解

$\|Au-b\|\le\|Ax-b\|,\ \forall x\in C^n$

$\|x_0\|_2\le\|u\|_2$

- 一、最佳最小二乘解

2、 Ax = b的最佳最小二乘解的计算

定理4.17 $x_0=A^+b$ 是 $Ax=b$ 的最佳最小二乘解。

证明：(1) 利用AA⁺：x_0是最小二乘解。由Th4.15

$C^m=R(AA^+)\oplus N(AA^+)=R(A)\oplus N(A^+),\ N(A^+)=R^\perp(A)$

$\Rightarrow \|AA^+b-b\|\le\|y-b\|,\forall y\in R(AA^+)=R(A)$

$\Rightarrow \|Ax_0-b\|\le\|Ax-b\|,\forall x\in C^n$

(2) 对任一最小二乘解u有：$u-x_0\in N(A)$，从而 $x_0\perp(u-x_0)$，

进而 $\|u\|^2=\|u-x_0+x_0\|^2=\|u-x_0\|^2+\|x_0\|^2\ge\|x_0\|^2$

设$b=b_1+b_2,\ b_1\in R(A),\ b_2\in N(A^+)=R^\perp(A)$, 则有

$\|Ax-b\|^2=\|Ax-b_1-b_2\|^2=\|Ax-b_1\|^2+\|b_2\|^2,\forall x\in C^n$

191

4.4 最佳最小二乘解

$\|Au-b\|\le\|Ax-b\|,\ \forall x\in C^n$

$\|x_0\|_2\le\|u\|_2$

- 一、最佳最小二乘解

2、 Ax = b的最佳最小二乘解的计算

定理4.17 $x_0=A^+b$ 是 $Ax=b$ 的最佳最小二乘解。

证明：(1) 利用AA⁺：x_0是最小二乘解：$\|Ax_0-b\|\le\|Ax-b\|,\forall x\in C^n$

(2) 对任一最小二乘解u有：$u-x_0\in N(A)$，从而 $x_0\perp(u-x_0)$，

进而 $\|u\|^2=\|u-x_0+x_0\|^2=\|u-x_0\|^2+\|x_0\|^2\ge\|x_0\|^2$

设$b=b_1+b_2,\ b_1\in R(A),\ b_2\in N(A^+)=R^\perp(A)$, 则有

$\|Ax-b\|^2=\|Ax-b_1-b_2\|^2=\|Ax-b_1\|^2+\|b_2\|^2,\forall x\in C^n$

$\Rightarrow\|Au-b_1\|\le\|Ax-b_1\|,\forall x\in C^n\ \Rightarrow\|Au-b_1\|=0\Rightarrow Au=b_1,Ax_0=b_1$

$\Rightarrow A(u-x_0)=b_1-b_1=0\quad\Rightarrow u-x_0\in N(A)$

$\because x_0\in R(A^+)=N^\perp(A),\ \therefore x_0\perp(u-x_0).$

192

例题1（*P. 106，eg8*）求不相容方程组Ax = b的最佳最小二乘解：$x_0=A^+b$。

- 例题2、设 $A=\begin{bmatrix}1&1&2\\0&2&2\\1&0&1\end{bmatrix}$，$\alpha=\begin{bmatrix}1\\1\\2\end{bmatrix}$

1. 证明：$\alpha\notin R(A)$
2. 在列空间R(A)上找一点y_0，y_0距离α最近。

$\alpha\notin R(A)\ \Leftrightarrow\ Ax=\alpha$ 无解 $(A\vdots\alpha)\to\begin{bmatrix}1&1&2&1\\0&2&2&1/2\\0&0&0&3/2\end{bmatrix}$

求 $x_0=A^+\alpha,\ y_0=Ax_0$

193

二、最佳拟合曲线

- 问题：在实际问题中，已知变量X和变量Y之间存在函数关系Y = F(X)，但不知道F(X)的具体形式，由观察和实验数据寻求经验公式：

  Y = f(X)，使得误差最小。
- 例题1（*P.107，eg9*）一组实验数据（1，2），（2，3），（3，5），（4，7）的分布呈直线趋势，求最佳拟合直线。
- 方法：
  - 确定系数β_1与β_2，使数据（x_i，y_i）最佳拟合$f=\beta_1+\beta_2 x$
  - 将误差向量表示为 $e=A\beta-b$，求方程组

    $A\beta-b=0$ 的最小二乘解β，由β给出拟合参数。

194

- 例题2（*P.107，eg10*）二次曲线拟合

  $\beta_1x^2+\beta_2y^2=1$
- 例题3

求参数 β，使得函数 $y=\beta_1x+\beta_2x^2+\beta_3\sin x$ 最佳拟合4个数据点：

$$(0,1),\ (\frac{\pi}{4},2),\ (\frac{\pi}{2},3),\ (\pi,4)$$

195

第5章、 矩阵分析

- 讨论：矩阵函数的分析性质
  - 函数的定义
  - 函数的计算
  - 函数的分析性质：连续、微分、积分等
- 定义矩阵函数的思想：
  - 用幂级数定义矩阵函数
  - 需要的背景概念：幂级数序列与收敛性质
- 本章的结构
  - 向量范数与矩阵范数（更一般的度量）
  - 向量序列和矩阵序列的收敛
  - 矩阵幂级数
  - 矩阵函数

Jordan化方法

196

§ 5.1 向量的范数

一、向量范数的概念（长度的推广）

1、赋范空间

可为一般空间V(F)

- 定义5.1（*P. 109*）$V_n(F)$上的实值函数||•||：

  $V_n(F)\to R^+$，满足
  1. 正定性：$\|x\|\ge 0$，$\|x\|=0\Leftrightarrow x=0$。
  2. 齐次性： $\forall k\in F$，$\|kx\|=|k|\,\|x\|$
  3. 三角不等式： $\|x+y\|\le\|x\|+\|y\|$

  则称 ||•|| 为$V_n(F)$上的范数，[$V_n(F)$；||•||] 是赋范空间。

2、C^n空间常用的范数 $\|x\|_p=\left(\sum_{i=1}^n|x_i|^p\right)^{\frac{1}{p}}$

- C^n空间，Hölder 范数（p- 范数）

197

- p-范数的特例（$p=2$: 长度）：
  - $p=1$ $\|x\|_1=\sum_{i=1}^n|x_i|$
  - $p=2$ $\|x\|_2=\sqrt{\sum_{i=1}^n|x_i|^2}$
  - $p=\infty$ $\|x\|_\infty=\max\{|x_i|,1\le i\le n\}$

范数不等式及相关概念： $\|x+y\|\le\|x\|+\|y\|$

$\|x-y\|\ge|\,\|x\|-\|y\|\,|$； $\|x-y\|+\|y\|\ge\|x\|$

$\|x+y\|\ge|\,\|x\|-\|y\|\,|$ $\|y-x\|+\|x\|\ge\|y\|$

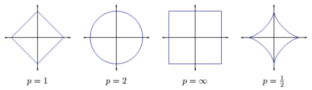
距 离：$d(x,y)=\|x-y\|$

邻域：$R(x_0,r)=\{x\in V_n(F),\|x-x_0\|<r\},\ r>0$

198

邻域：$R(x_0, r) = \{ x \in V_n(F), \|x - x_0\| < r \}, r > 0$

p-范数意义下，R^2中的单位球（邻域）



p = 1/2 为拟范数。

199

二、向量范数的收敛性质

- 1、向量范数的连续性

定理5.1 (*P. 110*) 在给定的赋范空间 $[V_n(F); \|\bullet\|]$ 上， $\{\alpha_1, \alpha_2, \ldots \alpha_n\}$是一组基，$\alpha=\Sigma a_i \alpha_i$，$\beta=\Sigma b_i \alpha_i$，则 $\forall \varepsilon >0$，$\exists \delta > 0$, 若有$\max_i |a_i - b_i|<\delta$，则有 $| \|\alpha\| - \|\beta\| | < \varepsilon$ 。

含义：$\| a-b \|_\infty \to 0 \Rightarrow | \|\alpha\| - \|\beta\| | \to 0$

向量范数是坐标的连续函数

2、向量范数的等价性

i. 等价的概念：$r_1 \|\alpha\|^{(1)} \le \|\alpha\|^{(2)} \le r_2 \|\alpha\|^{(1)}, \forall \alpha$

ii. 等价性定理（定理5.2 *P. 111* ）$V_n(F)$上的任意两种范数是等价的。

含义：$\|\alpha\|^{(1)} \to 0 \Leftrightarrow \|\alpha\|^{(2)} \to 0$

200

§ 5.2 矩阵的范数

- 一、矩阵范数
- 定义5.3（*P. 112*）$F^{n\times n}$上的实值函数$\|\bullet\|$：$F^{n\times n} \to R^+$，满足：$\forall A$

1. 正定性：$\| A \| \ge 0$，$\| A \| = 0 \Leftrightarrow A = 0$。
2. 齐次性： $\forall k \in F$，$\| kA \| = | k | \| A \|$
3. 三角不等式： $\| A+B \| \le \| A \| + \| B \|$
4. 相容性：$\| AB \| \le \| A \| \| B \|$

则称$\| A \|$为矩阵范数。

向量范数要“移植”成为矩阵范数，需满足相容性！

201

§ 5.2 矩阵的范数

例题1 A作为向量， 类似向量2-范数称为F-范数：

$$\|A\|_F = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^2 \right)^{\frac{1}{2}} = tr(A^H A)^{\frac{1}{2}}$$

例题2 设方阵A的奇异值为 $\sigma_1, \sigma_2, \cdots \sigma_r$，

则 $\|A\|_F^2 = tr(A^H A) = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2$.

例题3 设A有奇异值展开 $A = \sigma_1 u_1 v_1^H + \cdots + \sigma_r u_r v_r^H$，

$A_k = \sigma_1 u_1 v_1^H + \cdots + \sigma_k u_k v_k^H$， 则

$$\|A - A_k\|_F = \min \{ \|A - S\|_F, r(S) \le k \}.$$

202

例题4、证明对任何矩阵范数$\| A \|$，

1. $\| I \| \ge 1$
2. $\| A^n \| \le \| A \|^n$
3. A可逆，$\| A^{-1} \| \ge \| A \|^{-1}$

例题5 设矩阵A酉相似与B，则 $\| A \|_F = \| B \|_F$

- 二、诱导范数(向量范数诱导出的矩阵范数)
- 1、矩阵范数与向量范数的相容性
  - 定义5.4（*P. 114*）：$\| Ax \| \le \| A \| \| x \|$
  - 定理5.3 设$\| x \|$是向量范数，则下式是与之相容的矩阵范数，称为由向量范数$\| x \|$诱导的矩阵范数：$\|A\| = \max\limits_{x \neq 0} \left\{ \frac{\|Ax\|}{\|x\|} \right\}$

203

- 例题1 *p*-范数诱导的矩阵范数：

$$\|A\|_1 = \max_j \left\{ \sum_{i=1}^{n} |a_{ij}| \right\}$$ 最大列和范数

$$\|A\|_\infty = \max_i \left\{ \sum_{j=1}^{n} |a_{ij}| \right\}$$ 最大行和范数

$$\|A\|_2 = \sqrt{\lambda_{\max}} = \sigma_1$$ 谱范数

λ_{max}是$A^H A$的最大特征值。

例题2 设 $A = \begin{bmatrix} -1 & 0 & 2+i \\ 3 & 5 & 0 \\ 1 & 2 & -1 \end{bmatrix}$，求$\| A \|_1$

$= \max\{5, 7, \sqrt{5}+1\} = 7$

例题3 设U为n阶酉矩阵，则

$\|U\|_2 = 1$； $\|UA\|_2 = \|U^H A U\|_2 = \|A\|_2$ 。

204

- 例题4 设$A_{m\times n}$为给定的列满秩矩阵，$\| \bullet \|_\alpha$ 为F^m上的一种向量范数，则 $\|x\|_\beta = \|Ax\|_\alpha$为$F^n$上的一种向量范数。

例题5 设$\|\cdot\|$为$F^{n\times n}$上的一种矩阵范数，在F^n上定义，$\|x\|_\alpha = \|xE^T\|$，则 $\|x\|_\alpha$是F^n上的一种与矩阵范数$\|\cdot\|$相容的向量范数，其中E为F^n中分量均为1的向量。

205

§ 5. 3 向量序列和矩阵序列的极限

C^n中向量（$C^{n\times n}$中矩阵）序列的收敛性

1. 按分量（元素）收敛

向量序列 $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})^T \in C^n, k = 1, 2, \cdots$

按分量收敛 $\Leftrightarrow \{x_1^{(k)}\}, \{x_2^{(k)}\}, \cdots, \{x_n^{(k)}\}$ 均收敛，即

$$\lim_{k\to\infty} x_i^{(k)} = a_i, 1 \le i \le n.$$

记为 $\lim\limits_{k\to\infty} x^{(k)} = a = (a_1, a_2, \cdots, a_n)^T$；

矩阵序列 $A^{(k)} = (a_{ij}^{(k)}) \in C^{m\times n}, k = 1, 2, \cdots$

按元素收敛 $\Leftrightarrow$ $\{a_{ij}^{(k)}\}$ 均收敛 $(1 \le i, j \le n)$， 即

$\lim\limits_{k\to\infty} a_{ij}^{(k)} = a_{ij}, 1 \le i, j \le n.$ 记为 $\lim\limits_{k\to\infty} A^{(k)} = A = (a_{ij}).$

206

§ 5. 3 向量序列和矩阵序列的极限

C^n中向量（$C^{n\times n}$中矩阵）序列的收敛性

2. 按范数收敛:

$\{x^{(k)}\}$按范数$\|x\|$收敛于a $\Leftrightarrow \lim\limits_{k\to\infty} \|x^{(k)} - a\| = 0$;

$\{A^{(k)}\}$按范数$\|A\|$收敛于A $\Leftrightarrow \lim\limits_{k\to\infty} \|A^{(k)} - A\| = 0$.

3. 按分量（元素）收敛和按范数收敛的关系:

定理5.4（*P. 115*） C^n中一个向量序列$\{x^{(k)}\}$按分量收敛 $\Leftrightarrow$ 它按任意给定的范数收敛。

定理5.5（*P. 118*） $C^{n\times n}$中一个矩阵序列$\{A^{(k)}\}$按元素收敛 $\Leftrightarrow$ 它按任意给定的范数收敛。

207

4. 收敛向量或矩阵序列的性质：“四则”运算。

设 $a,b\in F, A^{(k)}_{n\times n}\to A, B^{(k)}_{n\times n}\to B$, 则

$$aA^{(k)}+bB^{(k)}\to aA+bB,\quad A^{(k)}B^{(k)}\to AB,$$
$$\det(A^{(k)})\to\det(A),\quad \|A^{(k)}\|\to\|A\|$$

若 $A^{(k)}$ ($k>>1$)与A均可逆，则 $(A^{(k)})^{-1}\to A^{-1}$.

例题 设方阵A构成的矩阵序列：

{I，A，A^2，…，A^k，…}

如果对某一矩阵范数|| A ||，有|| A || < 1，证明

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

由 $\|A^k\|\le\|A\|^k$ 可得所证。

5. 有界性：按分量（元素），或按范数

208

§ 5.4 矩阵的幂级数

讨论：方阵A和数列$\{a_i\}$构成的矩阵幂级数：

$$a_0I + a_1A + a_2A^2 + \dots + a_kA^k + \dots$$

1. 收敛性: 部分和序列敛散性
2. 求和方法，由和矩阵作为函数值定义矩阵函数。

一、谱半径

1. 定义：设A的谱为{λ_1，λ_2，…，λ_s}，则谱半径 $\rho(A) = \max\{|\lambda_i|,\ 1\le i\le s\}$。

例 $\rho(A^k) = (\rho(A))^k$，$\rho(kA) = |k|\rho(A)$，$\rho(A) = \rho(A^T)$；
相似矩阵的谱半径相同；
如果A是正规矩阵，则 $\rho(A) = \|A\|_2$。

定理3.13(P.82)：正规矩阵A的奇异值等于A的(非零)特征值的模

$A^k\to0 \Leftrightarrow \rho(A)<1.$ Jordan化方法！

209

§ 5.4 矩阵的幂级数

2 谱半径的性质：

1. 定理5.6：$A\in C^{n\times n}$，∀ 范数$\|A\|$，有 $\rho(A)\le\|A\|$。
2. 定理5.7 ∀ε>0，∃$\|A\|^*$，$\|A\|^*\le\rho(A)+\varepsilon$

含义：谱半径是任何矩阵范数的下确界。（下界中最大的）

数值范围（Th5.8, 5.9） 略

210

二、矩阵的幂级数

1. 定义及其收敛性：

矩阵幂级数：$a_0I+a_1A+a_2A^2+\dots+a_kA^k+\dots$

前n项部分和构成矩阵多项式序列 $S_n(A)=\sum_{k=0}^n a_kA^k$

$\sum_{k=0}^\infty a_kA^k$ 收敛 $\Leftrightarrow$ $\{S_n(A)\}$ 收敛

$\lim_{n\to\infty}S_n(A)=S \Rightarrow \sum_{k=0}^\infty a_kA^k=S$

幂级数的基本性质

设复数项幂级数 $f(z)=\sum_{k=0}^\infty a_kz^k$ 的收敛半径为R，则

$\sum_{k=0}^\infty a_k\frac{d^nz^k}{dz^n}$ $(=\frac{d^nf(z)}{dz^n})$ 在 $|z|<R$内收敛。

211

二、矩阵的幂级数

2. 收敛性的判别方法

定理5.10 设复数项幂级数 $\sum_{k=0}^\infty a_kz^k$ 的收敛半径为R。

(1) 若 $\rho(A)<R$，则 $\sum_{k=0}^\infty a_kA^k$ 收敛；

(2) 若 $\rho(A)>R$，则 $\sum_{k=0}^\infty a_kA^k$ 发散。

推论：若$\|A\|<R$，则 $\sum_{k=0}^\infty a_kA^k$ 收敛。

定理5.6 $\rho(A)\le\|A\|$

212

例题1、（P.108 eg8）讨论 $\sum_{k=0}^\infty A^k$ 的收敛性，在收敛时求和矩阵。

例题2、设 $A=\begin{bmatrix}0.2&0.1&0.2\\0.5&0.5&0.4\\0.1&0.3&0.2\end{bmatrix}$，讨论 $\sum_{k=0}^\infty A^k$ 的收敛性。

例题3、证明 级数 $\sum_{k=0}^\infty\frac{k^2}{10^k}\begin{bmatrix}1&2\\8&1\end{bmatrix}^k$ 收敛。

例题4 设 $A=\begin{bmatrix}-1&2&2\\0&-1&-1\\0&0&0\end{bmatrix}$，讨论 $\sum_{n=1}^\infty\frac{1}{n^2}A^n$ 的收敛性。

213

收敛半径R的求法 对 $\sum_{k=0}^\infty a_kz^k$ 有：

$$R=\frac{1}{\rho}$$

比值法 $\rho=\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right|$

根值法 $\rho=\lim_{k\to\infty}\sqrt[k]{|a_k|}$

214

§ 5.5 矩阵函数

一、矩阵函数的定义和性质

1 定义5.14 (p125)

$f(z)=\sum_{k=0}^\infty a_kz^k$ 是解析函数，收敛半径为R，如果 $\rho(A)<R$，则 $\sum_{k=0}^\infty a_kA^k$ 有意义。

- 定义矩阵函数 $f(A)=\sum_{k=0}^\infty a_kA^k$
- 常见的矩阵函数

$$e^A=\sum_{k=0}^\infty\frac{1}{k!}A^k,\ \sin(A)=\sum_{k=0}^\infty\frac{(-1)^k}{(2k+1)!}A^{2k+1},\ \rho(A)<+\infty;$$
$$(I-A)^{-1}=\sum_{k=0}^\infty A^k,\ \ln(I+A)=\sum_{k=1}^\infty\frac{(-1)^{k-1}}{k}A^k,\ \rho(A)<1,\ \cdots$$

215

§ 5.5 矩阵函数

$$(I+A)^{-1}=\sum_{k=0}^\infty(-1)^kA^k,\ \ln(I-A)=-\sum_{k=1}^\infty\frac{1}{k}A^k,\ \rho(A)<1$$
$$A^{-1}=?,\ \ln(A)=?$$
$$A^{-1}=(I+A-I)^{-1}=\sum_{k=0}^\infty(-1)^k(A-I)^k,\ \rho(A-I)<1$$
$$\ln(A)=\ln(I+A-I)=\sum_{k=1}^\infty\frac{(-1)^{k-1}}{k}(A-I)^k,\ \rho(A-I)<1$$

- 函数 e^A 的若干性质

1. 若AB = BA，则 $e^Ae^B = e^Be^A = e^{A+B}$，
2. $e^0 = I$，$e^I = eI$，
3. $(e^A)^{-1} = e^{-A}$。

216

$$A=\begin{bmatrix}1&1\\0&0\end{bmatrix}, B=\begin{bmatrix}1&-1\\0&0\end{bmatrix}\Rightarrow A^k=A, B^k=B, k>1$$

$$A+B=2E_{11}, \Rightarrow (A+B)^k=2^k E_{11}, k>0$$

$$e^A=I+(e-1)A \qquad e^B=I+(e-1)B$$

$$e^{A+B}=I+(e^2-1)E_{11}$$

- 例题1 设A为反对称矩阵，证明e^A为正交矩阵。

- 例题2 设 $A=\begin{bmatrix}3/2&0&2\\1&-3/2&0\\0&0&1\end{bmatrix}$，讨论 $\ln A$ 是否有意义

  ρ (A-I) = 5/2 > 1

217

二、矩阵函数的计算

- 计算 $f(A)=\sum_{k=0}^{\infty}a_kA^k$

1、Jordan标准形方法Th5.11

- A=PJP^{-1}，f(A)=Pf(J)P^{-1};
- 计算矩阵序列：S_n(J)
- 按元素收敛求得(在Th5.10的证明中，令N趋于无穷大！)

$$f(J_i(\lambda_i))=\begin{bmatrix}f(\lambda_i)&f'(\lambda_i)&\cdots&\frac{f^{(n_i-1)}(\lambda_i)}{(n_i-1)!}\\&f(\lambda_i)&\ddots&\vdots\\&&\ddots&f'(\lambda_i)\\&&&f(\lambda_i)\end{bmatrix}_{n_i\times n_i}$$

$$|f(A)|=\prod_{i=1}^{n}f(\lambda_i)$$

- 如果 A=PJP^{-1}，则 f(A)=Pf(J)P^{-1}；如果λ_i为矩阵A的特征值，则 $f(\lambda_i)$ 是矩阵f(A)的特征值;
- 含参数 t 的函数 f(At)。可与$f(\lambda t)$对应，t为参数。

218

二、矩阵函数的计算

- 计算 $f(A)=\sum_{k=0}^{\infty}a_kA^k$

A=PJP^{-1}，f(A)=Pf(J)P^{-1};

$$f(J_i(\lambda_i))=\begin{bmatrix}f(\lambda_i)&f'(\lambda_i)&\cdots&\frac{f^{(n_i-1)}(\lambda_i)}{(n_i-1)!}\\&f(\lambda_i)&\ddots&\vdots\\&&\ddots&f'(\lambda_i)\\&&&f(\lambda_i)\end{bmatrix}_{n_i\times n_i}$$

例题1 $A=\begin{bmatrix}2&0&0\\1&1&1\\1&-1&3\end{bmatrix}$，计算$e^A$和sin(A)等。

$$J_A=\begin{bmatrix}2&1&0\\0&2&0\\0&0&2\end{bmatrix}, P=\begin{bmatrix}0&1&1\\1&0&0\\1&0&-1\end{bmatrix}, P^{-1}=\begin{bmatrix}0&1&0\\1&-1&1\\0&1&-1\end{bmatrix}; \qquad \rho(A)=2, \quad |f(A)|=f(2)^3$$

$$f(A)=Pf(J_A)P^{-1}=P\begin{bmatrix}f(J_1)&\\&f(J_2)\end{bmatrix}P^{-1}=P\begin{bmatrix}f(2)&f'(2)&0\\0&f(2)&0\\0&0&f(2)\end{bmatrix}P^{-1}$$

219

二、矩阵函数的计算

- 计算 $f(A)=\sum_{k=0}^{\infty}a_kA^k$

A=PJP^{-1}，f(A)=Pf(J)P^{-1};

$$f(J_i(\lambda_i))=\begin{bmatrix}f(\lambda_i)&f'(\lambda_i)&\cdots&\frac{f^{(n_i-1)}(\lambda_i)}{(n_i-1)!}\\&f(\lambda_i)&\ddots&\vdots\\&&\ddots&f'(\lambda_i)\\&&&f(\lambda_i)\end{bmatrix}_{n_i\times n_i}$$

例题1 $A=\begin{bmatrix}2&0&0\\1&1&1\\1&-1&3\end{bmatrix}$，计算$e^A$和cos(A)等。

$$f(A)=P\begin{bmatrix}f(2)&f'(2)&0\\0&f(2)&0\\0&0&f(2)\end{bmatrix}P^{-1}=\begin{bmatrix}f(2)&0&0\\f'(2)&f(2)-f'(2)&f'(2)\\f'(2)&-f'(2)&f(2)+f'(2)\end{bmatrix}$$

故取R>2具体解析函数f(z)，则可求得相应的f(A)。

220

二、矩阵函数的计算

例题1 $A=\begin{bmatrix}2&0&0\\1&1&1\\1&-1&3\end{bmatrix}$，计算$e^A$和cos(A)等。

$$f(A)=\begin{bmatrix}f(2)&0&0\\f'(2)&f(2)-f'(2)&f'(2)\\f'(2)&-f'(2)&f(2)+f'(2)\end{bmatrix}, \qquad \rho(A)=2, \quad |f(A)|=f(2)^3$$

故取R>2的具体解析函数f(z)，则可求得相应的f(A)。

$f(z)=e^z$: $e^A=\begin{bmatrix}e^2&0&0\\e^2&0&e^2\\e^2&-e^2&2e^2\end{bmatrix}$, $\quad |e^A|=(e^2)^3=e^6$

221

二、矩阵函数的计算

例题1 $A=\begin{bmatrix}2&0&0\\1&1&1\\1&-1&3\end{bmatrix}$，计算$e^A$和cos(A)等。

$$f(A)=\begin{bmatrix}f(2)&0&0\\f'(2)&f(2)-f'(2)&f'(2)\\f'(2)&-f'(2)&f(2)+f'(2)\end{bmatrix}, \qquad \rho(A)=2, \quad |f(A)|=f(2)^3$$

故取R>2的具体解析函数f(z)，则可求得相应的f(A)。

$f(z)=\cos(z)$: $\cos(A)=\begin{bmatrix}\cos 2&0&0\\-\sin 2&\cos 2+\sin 2&-\sin 2\\-\sin 2&\sin 2&\cos 2-\sin 2\end{bmatrix}$

$|\cos A|=(\cos 2)^3$

222

二、矩阵函数的计算

例题1 $A=\begin{bmatrix}2&0&0\\1&1&1\\1&-1&3\end{bmatrix}$，计算$e^A$和cos(A)等。

$$f(A)=\begin{bmatrix}f(2)&0&0\\f'(2)&f(2)-f'(2)&f'(2)\\f'(2)&-f'(2)&f(2)+f'(2)\end{bmatrix}, \qquad \rho(A)=2, \quad |f(A)|=f(2)^3$$

故取R>2的具体解析函数f(z)，则可求得相应的f(A)。

f(z) = ln(1+z) : ？ 不行！（R=1）

f(z) = ln(1+z/3) : ？ 行！（|z/3|<1, R=3）

223

2、 最小多项式方法

定理5.12 设n阶方阵A的最小多项式为

$$m_A(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}, \quad \sum_{i=1}^{s}n_i=m$$

令 $g(\lambda)=c_0+c_1\lambda+c_2\lambda^2+\cdots+c_{m-1}\lambda^{m-1}$

$f(\lambda)$为解析函数，则 f(A) = g(A)

$$\Leftrightarrow g^{(j)}(\lambda_i)=f^{(j)}(\lambda_i), j=0,1,2,\cdots,n_i-1; i=1,2,\cdots,s$$

也可以用特征多项式代替最小多项式！

例题2 （P129 eg14）用法2计算上例

$J_A=\begin{bmatrix}2&1&0\\0&2&0\\0&0&2\end{bmatrix}$ $\Rightarrow m_A(\lambda)=(\lambda-2)^2 \Rightarrow m=2, g(\lambda)=c_0+c_1\lambda.$

由 $g(2)=f(2), g'(2)=f'(2)$ 得

$c_0=f(2)-2f'(2), c_1=f'(2)$

224

2、 最小多项式方法

定理5.12 设n阶方阵A的最小多项式为

$$m_A(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}, \quad \sum_{i=1}^{s}n_i=m$$

令 $g(\lambda)=c_0+c_1\lambda+c_2\lambda^2+\cdots+c_{k-1}\lambda^{m-1}$

$f(\lambda)$为解析函数，则 f(A) = g(A)

$$\Leftrightarrow g^{(j)}(\lambda_i)=f^{(j)}(\lambda_i), j=0,1,2,\cdots,n_i-1; i=1,2,\cdots,s$$

也可以用特征多项式代替最小多项式！

例题2 （P129 eg14）用法2计算上例

$J_A=\begin{bmatrix}2&1&0\\0&2&0\\0&0&2\end{bmatrix}$ $g(\lambda)=c_0+c_1\lambda, \ c_0=f(2)-2f'(2), c_1=f'(2)$

$\Rightarrow f(A)=c_0I+c_1A$ 与例1相同！

225

2、 最小多项式方法

例题3 （P129 eg15）计算 e^{At} $\sim f(\lambda)= e^{\lambda t}$

$$A=\begin{bmatrix}0&0&-2\\0&1&0\\1&0&3\end{bmatrix},\quad |\lambda I-A|=(\lambda-1)^2(\lambda-2),$$

$$m_A(\lambda)=(\lambda-1)(\lambda-2),\ \Rightarrow g(\lambda)=c_0+c_1\lambda.$$

由 $g(1)=f(1),g(2)=f(2),$

即 $g(1)=c_0+c_1=f(1)=e^t, g(2)=c_0+2c_1=f(2)=e^{2t},$

$$\Rightarrow c_1=e^{2t}-e^t, c_0=2e^t-e^{2t},\quad \Rightarrow e^{At}=c_0I+c_1A.$$

例题4 设 $A=\begin{bmatrix}1&1&0\\0&0&1\\0&0&1\end{bmatrix}$，计算$A^{10}$。$\sim f(\lambda)= \lambda^{10}$

226

§ 5.6 函数矩阵的微积分

一、函数矩阵及其分析性质

- 函数矩阵：$A(t)=[a_{ij}(t)]_{m\times n}$，
- 分析性质：
- A(t) 连续、可微分、可积分 ⇔ $a_{ij}(t)$ {连续 / 可微分 / 可积分}

$$\lim_{t\to t_0}A(t)=[\lim_{t\to t_0}a_{ij}(t)]_{m\times n}$$

$$\frac{dA(t)}{dt}=[\frac{da_{ij}(t)}{dt}]_{m\times n}$$

$$\int_a^b A(t)dt=[\int_a^b a_{ij}(t)dt]_{m\times n}$$

- 微分性质 (p130)

227

- 微分性质 (p130)

$$(aA(t)+bB(t))'=aA'(t)+bB'(t);$$

$$(A(t)B(t))'=A'(t)B(t)+A(t)B'(t);$$

$$(A^{-1}(t))'=-A^{-1}(t)A'(t)A^{-1}(t);$$

$$(A^2(t))'=A'(t)A(t)+A(t)A'(t);$$

- 初等函数的微分性质

$$(e^{At})'=Ae^{At}=e^{At}A;$$

$$(\sin At)'=A\cos At=(\cos At)A;(\cos At)'=-A\sin At=-(\sin At)A;$$

$$\ln(I+At)'=A(I+At)^{-1}=(I+At)^{-1}A, \rho(At)=|t|\rho(A)<1;\cdots$$

228

- 积分性质 (p131)

$$\int(aA(t)+bB(t))dt=a\int A(t)dt+b\int B(t)dt;$$

$$\int CA(t)dt=C\int A(t)dt;$$

$$\int A(t)B'(t)dt=A(t)B(t)-\int A'(t)B(t)dt;$$

- 原函数与积分计算 设 $A'(t)=B(t)$

则 $\int B(t)dt=A(t)+C;\quad \int_a^b B(t)dt=A(b)-A(a);\cdots$

如 $\int Ae^{At}dt=e^{At}+C;$

$$\int A\cos(At)dt=\sin At+C;\ \int A\sin(At)dt=-\cos At+C;\cdots$$

229

§ 5.7 矩阵函数的应用（求解常系数微分方程组）

一、微分方程组的一般形式

$X'(t)=A(t)X(t)+f(t)$

$X(t_0)=C$。

{齐次：$f(t)=0$ / 非齐次：$f(t)\neq 0$ / 常系数：$A(t)=A$}

二、一阶线性常系数齐次微分方程组

求解：$X'(t)=AX(t)$， $X(t_0)=C$。

定理5.13 上述方程组的解为：

$$X(t)=e^{A(t-t_0)}X(t_0)$$

例题1 求解 $X'(t)=\begin{bmatrix}1&2\\4&3\end{bmatrix}X(t),\ X(0)=\begin{bmatrix}1\\2\end{bmatrix}$

230

三、一阶线性常系数非齐次微分方程组

求解：$X'(t)=AX(t)+f(t)$，$X(t_0)=C$。

定理5.14(P133)：上述方程组的解为：

$$X(t)=e^{A(t-t_0)}X(t_0)+\int_{t_0}^{t}e^{A(t-s)}f(s)ds$$

证 由 $(e^{-At}X(t))'=-Ae^{-At}X(t)+e^{-At}X'(t)=e^{-At}f(t)$

两边积分，即得所求。

例题2 求解 $X'(t)=\begin{bmatrix}1&2\\4&3\end{bmatrix}X(t)+\begin{bmatrix}1\\-1\end{bmatrix},\ X(0)=\begin{bmatrix}1\\2\end{bmatrix}$

$$\det(\lambda I-A)=(\lambda+1)(\lambda-5)$$
$$\Rightarrow m_A(\lambda)=(\lambda+1)(\lambda-5)$$
$$\Rightarrow X(t)=e^{5t}\begin{bmatrix}1\\2\end{bmatrix}+(1-e^{-t})\begin{bmatrix}1\\-1\end{bmatrix}$$

231

讨论问题：

从矩阵函数计算中可以得到矩阵函数 $f(A)$ 的哪些有关信息？

1. 从Jordan标准形方法
   - $f(A)$的特征值
   - $f(A)$相似性
2. 从最小多项式方法
   - $f(A)=g(A)$：任何一个n阶方阵的矩阵函数可以用一个次数不超过n的矩阵多项式来表示。
   - $A f(A)=f(A)A$：任何方阵和它的矩阵函数乘法可交换。

232

第 6 章
矩阵的Kronecker积和Hadamard积

The Kronecker Product and Hadamard Product

233

概述：

主要内容：

- 介绍Kronecker积和Hadamard积
- 讨论
  - K-积，H-积的运算性质、之间的关系
  - K-积与矩阵乘积的关系
  - K-积，H-积的矩阵性质
  - K-积的矩阵等价与相似关系
- 应用：求解矩阵方程
  - 向量化算子
- 重点：K-积及其应用

234

6.1 Kronecker积和Hadamard积的定义

- 定义6.1（P. 136）
  - 设矩阵 $A=[a_{ij}]_{m\times n}$和 $B=[b_{ij}]_{s\times t}$，则A和B的Kronecker被定义为A⊗B:

$$A\otimes B=[a_{ij}B]_{ms\times nt}=\begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

设 $A=[a_{ij}]_{m\times n}$和 $B=[b_{ij}]_{m\times n}$为同阶矩阵，则A和B的Hadamard被定义为A∘B:

$$A\circ B=[a_{ij}b_{ij}]_{m\times n}=\begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1n}b_{1n} \\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2n}b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1}b_{m1} & a_{m2}b_{m2} & \cdots & a_{mn}b_{mn} \end{bmatrix}$$

235

6.1 K-积和H-积的定义

$A\otimes B=[a_{ij}B]$

$A\circ B=[a_{ij}b_{ij}]$

例题1 设 $A=\begin{bmatrix}1 & 3\\ 2 & 4\end{bmatrix}$, $B=\begin{bmatrix}3 & 0\\ 0 & -1\end{bmatrix}$，计算

A⊗B，B⊗A，I_2⊗B，A∘B，I_2∘A

$$A\otimes B=\begin{bmatrix} 1\times\begin{bmatrix}3 & 0\\ 0 & -1\end{bmatrix} & 3\times\begin{bmatrix}3 & 0\\ 0 & -1\end{bmatrix} \\ 2\times\begin{bmatrix}3 & 0\\ 0 & -1\end{bmatrix} & 4\times\begin{bmatrix}3 & 0\\ 0 & -1\end{bmatrix}\end{bmatrix}=\begin{bmatrix}3 & 0 & 9 & 0\\ 0 & -1 & 0 & -3\\ 6 & 0 & 12 & 0\\ 0 & -2 & 0 & -4\end{bmatrix},$$

$A\otimes B$

$\neq B\otimes A$

$$B\otimes A=\begin{bmatrix} 3\times\begin{bmatrix}1 & 3\\ 2 & 4\end{bmatrix} & 0\times\begin{bmatrix}1 & 3\\ 2 & 4\end{bmatrix} \\ 0\times\begin{bmatrix}1 & 3\\ 2 & 4\end{bmatrix} & -1\times\begin{bmatrix}1 & 3\\ 2 & 4\end{bmatrix}\end{bmatrix}=\begin{bmatrix}3 & 9 & 0 & 0\\ 6 & 12 & 0 & 0\\ 0 & 0 & -1 & -3\\ 0 & 0 & -2 & -4\end{bmatrix},$$

236

6.1 K-积和H-积的定义

$A\otimes B=[a_{ij}B]$

$A\circ B=[a_{ij}b_{ij}]$

例题1 设 $A=\begin{bmatrix}1 & 3\\ 2 & 4\end{bmatrix}$, $B=\begin{bmatrix}3 & 0\\ 0 & -1\end{bmatrix}$，计算

A⊗B，B⊗A，I_2⊗B，A∘B，I_2∘A

$$I_2\otimes B=\begin{bmatrix}B & 0\\ 0 & B\end{bmatrix}=\begin{bmatrix}3 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & 3 & 0\\ 0 & 0 & 0 & -1\end{bmatrix},$$ 分块对角矩阵

$$A\circ B=\begin{bmatrix}1\times 3 & 3\times 0\\ 2\times 0 & 4\times(-1)\end{bmatrix}=\begin{bmatrix}3 & 0\\ 0 & -4\end{bmatrix},\quad A\circ B=B\circ A$$

$$I_2\circ A=\begin{bmatrix}1\times 1 & 0\times 1\\ 0\times 2 & 1\times 4\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 4\end{bmatrix}.$$ 对角矩阵

237

6.1 K-积和H-积的定义

例题2 设分块矩阵A = (A_{st})，则

A⊗B = (A_{st} ⊗ B)

特别地，若A = (A_1, A_2, …, A_n)，则

A⊗B = (A_1⊗B, A_2⊗B,…, A_n⊗B)

例题3 快速Walsh(Hadamard)变换 $y_N = H_N x_N$，

其中

$$H_N=\begin{bmatrix}H_{N/2} & H_{N/2}\\ H_{N/2} & -H_{N/2}\end{bmatrix}, N=2^n, n=1,2,\cdots, H_1=[1].$$

于是有

$$H_N=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\otimes H_{N/2}=H_2\otimes H_{N/2}=\cdots=H_2^{\otimes n}.$$

$$H_N=\begin{bmatrix}H_{N/2} & \\ & H_{N/2}\end{bmatrix}\begin{bmatrix}I_{N/2} & I_{N/2}\\ I_{N/2} & -I_{N/2}\end{bmatrix}=(I_2\otimes H_{N/2})(H_2\otimes I_{N/2})$$

238

6.1 K-积和H-积的定义

例题2 设分块矩阵A = (A_{st})，则

A⊗B = (A_{st} ⊗ B)

特别地，若A = (A_1, A_2, …, A_n)，则

A⊗B = (A_1⊗B, A_2⊗B,…, A_n⊗B)

例题3 快速Walsh(Hadamard)变换 $y_N = H_N x_N$，

其中

$$H_N=\begin{bmatrix}H_{N/2} & H_{N/2}\\ H_{N/2} & -H_{N/2}\end{bmatrix}, N=2^n, n=1,2,\cdots, H_1=[1].$$

于是有

$$H_N=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\otimes H_{N/2}=H_2\otimes H_{N/2}=\cdots=H_2^{\otimes n}.$$

$$H_N=\begin{bmatrix}I_{N/2} & I_{N/2}\\ I_{N/2} & -I_{N/2}\end{bmatrix}\begin{bmatrix}H_{N/2} & \\ & H_{N/2}\end{bmatrix}=(H_2\otimes I_{N/2})(I_2\otimes H_{N/2})$$

239

- K-积，H-积的基本结果：
  - A和B中有一个为零矩阵，则 A⊗B=0，A∘B=0
  - I⊗I=I，I∘I=I
  - 若A为对角矩阵，则A⊗B为分块对角矩阵，A∘B为对角矩阵。
- K-积的基本性质
  - 定理6.1（P. 138）设以下矩阵使计算有意义，则
    - (kA)⊗B = A⊗(kB)
    - A⊗(B + C) = A⊗B + A⊗C
    - (A⊗B)⊗C = A⊗(B⊗C)
    - $(A\otimes B)^H = A^H \otimes B^H$
    - A⊗B ≠ B⊗A

240

- H-积的基本性质：

  设A，B为同阶矩阵，则
  - A∘B = B∘A
  - (kA)∘B = A∘(kB)
  - A∘(B + C) = A∘B + A∘C
  - (A∘B)∘C = A∘(B∘C)
  - $(A\circ B)^H = A^H \circ B^H$
- Kronecker和Hadamard的关系：
  - 定理6.3（P. 139） A∘B 可由A⊗B的元素构成。

241

- K-积与矩阵乘法
  - 定理6.2（P. 138）设矩阵A，B，C，D使得下列运算有意义，则有

    (A⊗B) (C⊗D) = (AC) ⊗ (BD)
  - 意义：建立Kronecker积和矩阵乘法的相互转换。
  - 特别情形：设 $A\in F^{m\times m}$，$B \in F^{n\times n}$，则
  - $A\otimes B = (I_mA)\otimes(BI_n) = (AI_m)\otimes(I_nB)$

    $= (I_m\otimes B) (A\otimes I_n) = (A\otimes I_n) (I_m\otimes B)$
  - $(A\otimes B)^k = A^k \otimes B^k$

$(A_1\otimes B_1\otimes C_1)(A_2\otimes B_2\otimes C_2) = (A_1A_2)\otimes(B_1B_2)\otimes(C_1C_2)$

$(A_1\otimes B_1)(A_2\otimes B_2)(A_3\otimes B_3) = (A_1A_2A_3)\otimes(B_1B_2B_3)$

242

6.2 Kronecker积和Hadamard积的性质

- Kronecker积的矩阵性质
  - 定理6.4（P. 140）设矩阵使下列运算有意义，则
    - 当A，B分别为可逆矩阵时，A⊗B和B⊗A均为可逆矩阵，而且有 $(A\otimes B)^{-1} = A^{-1} \otimes B^{-1}$
    - 当方阵$A\in F^{m\times m}$，$B\in F^{n\times n}$时，方阵$A\otimes B\in F^{mn\times mn}$的行列式为 $|A\otimes B| = |B\otimes A| = |A|^n |B|^m$
    - 若A，B是Hermite矩阵，则A⊗B 和B⊗A均是Hermite矩阵
    - 若A，B是酉矩阵，则A⊗B和B⊗A均是酉矩阵。

243

Kronecker与矩阵等价、相似关系

定理6.5（P. 141）

- 设矩阵A，B，为等价矩阵，则(A⊗I)等价于(B⊗I)
- 设方阵A相似与J_A，方阵B相似于J_B，则(A⊗B) 相似于($J_A \otimes J_B$)

K-积特征值和特征向量

定理6.6（P . 142）设A∈$F^{m\times m}$的特征值、特征向量分别是λ_i，x_i，B ∈ $F^{n\times n}$的特征值、特征向量分别是 μ_j， y_j，则

- (A⊗B) 的特征值是$\lambda_i\mu_j$ 。特征向量是($x_i \otimes y_j$)。
- ($A\otimes I_n$) +($I_m \otimes B$) 的特征值是$\lambda_i+\mu_j$， 特征向量是($x_i \otimes y_j$)

Kronecker和，记为A⊕B

244

Kronecker与矩阵等价、相似关系

推论

- 若A，B正定（半正定），则A⊗B和A⊕B均正定（半正定）；
- 若A相似于J_A，B相似于J_B，则 A⊗B 相似于$J_A \otimes J_B$，A⊕B 相似于 $J_A \oplus J_B$。
- 更一般的结果：

定理6.7（P. 142） $P(A,B)=\sum_{i,j=0}^{T} c_{ij}A^i \otimes B^j$

的特征值为 $P(\lambda_r,\mu_t)=\sum_{i,j=0}^{T} c_{ij}\lambda_r^{\,i}\mu_t^{\,j}$

245

Kronecker积的矩阵函数性质

定理6.8（P. 143）设是$f(z)$解析函数，$f(A)$有意义，则

- $f(I\otimes A) = I\otimes f(A)$　　$S_N(I\otimes A) = I\otimes S_N(A)$
- $f(A\otimes I) = f(A)\otimes I$　　$S_N(A\otimes I) = S_N(A)\otimes I$

特例：

- $e^{I_m\otimes A}=I_m\otimes e^A$
- $e^{A\otimes I_m}=e^A\otimes I_m$

定理的证明思路：利用定理5.12，矩阵函数可由多项式表示。也可以直接用极限性质证明。

246

例题1 设 A∈$F^{m\times n}$， B∈$F^{s\times t}$， 证明

rank (A ⊗ B) = rank (A) rank (B)

例题2（P . 144），设

$$A=\begin{bmatrix}3&-1\\0&1\end{bmatrix},\ B=\begin{bmatrix}2&0\\1&-1\end{bmatrix}$$

- 求(A⊗B)的特征值和特征向量
- 求[(A⊗I) +(I⊗B)]的特征值和特征向量

例题3：证明对任何方阵A, B, 有

$$e^{A\oplus B}=e^A\otimes e^B=e^B\otimes e^A$$

247

Hadamard积的性质

定理6.9（Schur积定理）设A、B为同阶方阵。若A和B半正定（正定），则A∘B亦半正定（正定）。

证明思路：利用定理3.6，有

$$A=\sum_{r=1}^{k}v_rv_r^H,\ B=\sum_{s=1}^{l}w_sw_s^H,$$

推出 A∘B可表示为

$$A\circ B=\sum_{r=1}^{k}\sum_{s=1}^{l}u_{rs}u_{rs}^H,\ u_{rs}=v_r\circ w_s.$$

248

6. 3 矩阵的向量化算子和K-积

向量化算子Vec: $F^{m\times n}\to F^{mn}$

- 定义（P . 143）设 A = $[a_{ij}]_{m\times n}$，则

Vec(A) = ($a_{11}\, a_{21} \ldots a_{m1}$; $a_{12}\, a_{22} \ldots a_{m2}$; …; $a_{1n}\, a_{2n} \ldots a_{mn}$)^T

- 性质：（P. 146）

1. Vec是线性算子，并保持线性关系不变：
   Vec (k_1A+k_2B) = k_1Vec (A) + k_2Vec (B)
2. 定理6. 10（P. 146）Vec(ABC) = (C^T ⊗ A)VecB
3. Vec(AX) = (I ⊗A)VecX　　令 B = X, C = I
4. Vec(XC) = (C^T⊗I)VecX　　令 B = X, A = I

249

用向量化算子求解矩阵方程

思想：用Vec算子，结合Kronecker积将矩阵方程化为线性方程组求解。

1、A∈$F^{m\times m}$， B∈$F^{n\times n}$， D∈$F^{m\times n}$， AX+XB = D

分析：

AX + XB = D ⇔ (I⊗A + B^T⊗I)VecX = VecD

G = (I⊗A + B^T⊗I)，

方程有惟一解的充要条件是G为可逆矩阵，即A和-B没有共同的特征值。

例题1 （P. 147）

250

用向量化算子求解矩阵方程

2、A，X∈$F^{n\times n}$， AX – XA = kX

分析：

AX – XA = kX ⇔ (I ⊗A–A^T⊗I)VecX = kVecX

H = (I⊗A – A^T⊗I)，

方程 (kI – H)y = 0 有非零解的充要条件是k为H的特征值，$k = \lambda_i - \lambda_j$。

例题2 求解矩阵方程 AX – XA = –2X

$$A=\begin{bmatrix}1&0\\2&3\end{bmatrix}$$

251

用向量化算子求解矩阵方程

3　A，B，D，X∈$F^{n\times n}$， AXB = D

分析：AXB = D ⇔ (B^T⊗A)VecX = VecD

L = B^T ⊗ A ，

方程有惟一解的充要条件是L为可逆矩阵.

例题3 求解方程 $A_1XB_1 + A_2XB_2 = D$

$$A_1=\begin{bmatrix}2&2\\2&-1\end{bmatrix}\quad B_1=\begin{bmatrix}1&0\\-1&1\end{bmatrix}\quad A_2=\begin{bmatrix}0&1\\-2&-1\end{bmatrix}$$

$$B_2=\begin{bmatrix}0&2\\-1&2\end{bmatrix}\quad D=\begin{bmatrix}4&-6\\0&8\end{bmatrix}$$

252

例题4 设 $A\in C^{m\times m}$，$B\in C^{n\times n}$，$D\in F^{m\times n}$，证明谱半径 $\rho(A)\cdot\rho(B)<1$ 时方程：

$$X = AXB + D$$

的解为 $X=\sum_{k=0}^{\infty}A^{k}DB^{k}$

证 $(I-B^{T}\otimes A)Vec(X)=Vec(D)$

$$\Rightarrow Vec(X)=(I-B^{T}\otimes A)^{-1}Vec(D)$$
$$=\sum_{k=0}^{\infty}(B^{T}\otimes A)^{k}Vec(D)$$
$$=\sum_{k=0}^{\infty}Vec(A^{k}DB^{k})$$

253

用向量化算子求解矩阵微分方程

4 A，B，$X\in F^{n\times n}$，

$$X'(t)=AX(t)+X(t)B,\ X(t_0)=C$$

$$\Leftrightarrow VecX'(t)=(I\otimes A+B^{T}\otimes I)VecX(t),$$
$$VecX(t_0)=VecC。$$

254

交换矩阵 K_{mn} 及其性质

$$K_{mn}=\sum_{i=1}^{m}\sum_{j=1}^{n}E_{ij}\otimes E_{ij}^{T},\ E_{ij}\in F^{m\times n}$$

定理6.11 (1) $K_{mn}^{T}=K_{nm}$; (2) $K_{1n}=K_{n1}=I_n$;

(3) $K_{mn}=\sum_{j=1}^{n}e_j^{T}\otimes I_m\otimes e_j$.

定理6.12 设 $A=(a_{ij})_{m\times n}$ 则

$$Vec(A)=K_{mn}^{T}Vec(A^{T}).$$

定理6.13 设 $A\in F^{m\times p}$，$B\in F^{n\times q}$，则

$$K_{mn}(B\otimes A)K_{pq}^{T}=A\otimes B.$$

255

复习选讲：

- 线性空间的表示
- 线性变换与变换矩阵
  - 线性变换的确定方法
  - 相应变换矩阵的求法
- 矩阵分解与空间分解
  - 准对角矩阵分解与不变子空间的分解
  - 可对角化矩阵的分解与特征子空间的分解
  - 幂等矩阵的空间分解
- J_A，$m_A(\lambda)$，$f(\lambda)=|\lambda I-A|$ 之间的关系
- A与 $f(A)$ 在Jordan标准形上的关系
  - $f(A)$ 的矩阵性质

256

复习选讲

- 正规矩阵的性质与应用
- 向量范数与矩阵范数
  - 向量的p范数
  - 矩阵的F范数和p范数
- 矩阵幂级数和矩阵函数
  - 矩阵幂级数的收敛与矩阵函数的意义
  - 矩阵幂级数的求和与矩阵函数的计算
  - 矩阵函数与矩阵多项式

257

习题选讲

- P150：9
- P31：1(3), 17,
- P58：6, 11，20
- P92：11，12，15，

258