线性代数扫描笔记

Peize Liu*

Shenzhen Middle School

目录

*Email: peize.liu@spc.ox.ac.uk

前言

最开始自学线性代数的动机是看不懂电动力学的第一章，加上高三在学校选了线代这门课程，促使我抽出业余时间开始看国内外有关教材。高三因为申请季的原因，前前后后拖了快一年才结束第一轮的学习，并留下了一本 200 来页的手写笔记。由于学习时间跨度很大，随着理解的深入，笔记前后的风格出入明显；而且参考的教材略多，总想写得面面俱到，内容的安排稍显散乱。个人水平所限，错漏之处想必不少，还请谅解。

内容概要

内容基本覆盖了传统的理工科线性代数内容，以及与其他课程联系紧密的重线性代数（张量）。与国内数学系大一必修的高等代数不同的是，这个笔记缺少多项式理论和环与理想，这是因为这些内容完全可以放在接下来的抽象代数中学习；这个笔记没有深入讨论 Jordan 标准形理论，内容仅限于得到复向量空间上线性算子的准素分解与循环分解，而不涉及 λ 矩阵，有理标准形，矩阵函数等课题。请注意由于学习时间跨度长，笔记前后风格差异相当大。

分章节的内容简介：

第一章*线性方程组*介绍了线性方程组的行化简法，矩阵行变换，阶梯形矩阵，$\mathbb{R}^n$ 的向量，线性无关，张成空间，矩阵乘向量表示 $\mathbb{R}^n$ 中的线性映射。

第二章*矩阵*介绍了矩阵加法与数乘，矩阵乘法表示线性映射的复合，矩阵转置，矩阵求逆，初等矩阵，矩阵分块，对角矩阵，三角矩阵，准对角矩阵，矩阵 LU 分解。

第三章*行列式*介绍了行列式表示线性映射放大率，行列式的基本性质（多线性，斜对称性），行列式的导出性质，行列式的完全展开，余子式展开，Vandermonde 行列式，分块矩阵的行列式，Cramer 法则，伴随矩阵表示逆矩阵。

第四章*向量空间*介绍了向量空间的公理及导出性质，子空间，基，维数，线性映射，单射与满射，零空间与值域，线性映射基本定理，空间的同构，矩阵表示线性映射，线性算子，空间的坐标系，基变换，相似矩阵，矩阵的 4 个基本子空间，矩阵的秩，子空间的和，直和，维数定理，积空间，商空间，商映射。

第五章*谱理论*介绍了本征值，本征向量，本征空间，特征多项式，相似矩阵的行列式与迹，代数重数与几何重数，对角化的等价条件，对角化算法，Fibonacci 数列，不变子空间，上三角矩阵，算子的多项式，零化多项式，Cayley-Hamilton 定理，准素分解，极小多项式，幂零算子，循环子空间分解，幂零算子的 Jordan 形，一般算子的 Jordan 形，Jordan-Chevalley 分解定理，广义本征空间，向量空间和算子的复化。

第六章*二次形*介绍了双线性形式，合同矩阵，二次型与对称矩阵，对称矩阵的合同对角化，化二次型为标准形，二次型的规范形，Sylvester 惯性定律，实二次型的分类，Sylvester 正定性准则。

第七章*内积空间*介绍了 Euclid 空间与酉空间的内积定义，范数，Cauchy-Schwarz 不等式，勾股定理，正交分解，赋范空间，标准正交基，Fourier 分解，Parseval 等式，Gram-Schmidt 正交化过程，Schur 定理，Riesz 表示定理，正交补，正交投影，极小化，正规方程，伴随映射，自伴算子，实谱定理，实对称矩阵的正交对角化，正规算子，复谱定理，等距同构，正交矩阵与酉矩阵，正交算子的刻画，QR 分解，初等旋转，正算子，极分解，Schmidt 分解，奇异值分解，算子范数。

第八章*重线性代数*介绍了线性泛函，对偶空间，对偶映射，零化子，Euclid 空间中对偶映射与伴随映射的等同，多重线性映射，张量积的万有性，张量积的基与维数，张量积的交换律与结合律，线性映射的张量积，矩阵的 Kronecker 积，对偶的张量积，Einstein 约定记号，协变量与逆变量，(p, q) 型张量，张量的运算，张量的基变换，指标缩并，交错映射，外积，交错张量，交错化子，Grassmann 代数，行列式。

和通用符号的差异

由于个人习惯和 follow 的书本等原因，笔记有许多符号表示和通用记号有差异（但笔记前后尽可能保持一致）

笔记中符号 $\mathbb{F}$ 仅表示实数域 $\mathbb{R}$ 或复数域 $\mathbb{C}$ ，不扩展到一般域的情况。

如无特殊指明，向量空间 V 不默认为有限维空间；但一旦指明空间的基，就一定是有限维的情况（笔记中只定义了有限维向量空间的基，并且尽力避免无限求和的情况）。

所有向量我都按习惯加上了箭头，例如矩阵作用于向量 $AX=B$ 习惯写作 $A\vec{x}=\vec{b}$ ；双线性形式 $f=X^{\mathrm{T}}AY$ 习惯写作 $f=\vec{x}^{\mathrm{T}}A\vec{y}$ ；向量空间的基习惯写作 $\vec{v}_1,\vec{v}_2,...,\vec{v}_n$。这个习惯直到最后一章重线性代数才放弃（因为涉及非常多不同的向量空间）。

为作区分，线性映射一般用字母 S,T,U ，矩阵一般用字母 A,B,C 。线性映射的核空间 $\mathrm{Ker}\,T$ 习惯写作零空间 $\mathrm{Null}\,T$ ；像集 $\mathrm{Im}\,T$ 习惯写作值域 $\mathrm{Ran}\,T$ 。

方阵/算子的行列式 $|A|$ 习惯记作 $\det A$ ，例如特征多项式 $|\lambda E-A|$ 习惯记作 $\det(\lambda I-A)$ 。我一直不赞同竖线 $|\cdot|$ 的滥用，容易造成混淆。在笔记中 $|\cdot|$ 只用于表示绝对值，而范数则用双竖线 $||\cdot||$ 表示。

同样造成滥用的还有圆括号 $(\cdot,\cdot)$ ，笔记中内积改用尖括号 $\langle\cdot,\cdot\rangle$ 表示。

与伴随作区分，对偶空间 V^* 和对偶映射 T^* 习惯写作 V' 和 T' （这种区分事实上对 Euclid 空间是无必要的）。

许多名词的中译与国内教材有差别，例如“线性变换”都称为“算子”。笔记中的名词基本都是跟随 Linear Algebra Done Right 的中译本。

证明结尾放上中二的 $\mathscr{Q}.\mathscr{E}.\mathscr{D}.$!

参考教材

我参考或者浏览过不少教材，对我写作影响最大的莫过于

Sheldon Axler, *Linear Algebra Done Right (3rd ed.)*

这是我最为推崇的线性代数教材。我的逻辑主线很大程度上继承了这本书的思路，有几个章节（向量空间、内积空间）甚至是完全照搬原书脉络。

此外，我还在很多章节主要参考了两本书：

蓝以中, 高等代数简明教程
Sergei Treil, *Linear Algebra Done Wrong*
张贤科, 高等线性代数

在某一两个知识点上给以我启发的书有：

丘维声, 高等代数
Steven Roman, *Advanced Linear Algebra*
David Lay, *Linear Algebra and Its Applications*
Gilbert Strang, *Introduction to Linear Algebra*
Kenneth Hoffman & Ray Kunze, *Linear Algebra*
梁灿彬 & 周彬, 微分几何入门与广义相对论

线性方程组与向量

1.1 Gauss消元法 & 行化简法

- 含未知数 $x_1, x_2, \cdots, x_n$ 的线性方程：

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \ , \ a_1, a_2, \cdots, a_n, b \in F$$

- 线性方程组：

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2 \\ \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n = b_m \end{cases}$$

记号 F：表示一个数域 这里指 R 或 C，后同

未知量 x 前的系数写成系数矩阵：

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} = \{a_{ij}\} = A$$

这样一个 m 行 n 列的矩阵，称为 $m \times n$ 矩阵。

线性方程组的增广矩阵：

$$\left[\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} & b_m \end{array}\right]$$

- Gauss消元法（加减消元法）

$\Downarrow$

- 矩阵的行变换：
  ① 倍加变换：把某一行的倍数加到另一行上
  ② 对换变换：把两行交换
  ③ 倍乘变换：把某一行的所有元素同乘一个非零数

- 矩阵的行变换属于初等变换（=行变换+列变换）.
  初等变换是可逆的.
- 行等价：若一个矩阵经过若干次行变换可变为另一个矩阵，则两矩阵行等价，记为 $A \sim B$.
  若两个线性方程组的增广矩阵行等价，则它们的解集相同.
- 先导元素：非零行最左侧的非零元素.
- 阶梯形矩阵：
  ①每个非零行在每个零行之上
  ②每个先导元素在前一行先导元素的右侧.
- 简化阶梯形矩阵：
  ①阶梯形矩阵
  ②所有先导元素等于1
  ③每个先导元素都是该列的唯一非零元素.
  唯一性：每个矩阵行等价于唯一的简化阶梯形矩阵
- 主元：先导元素的更常见说法.
- 行化简法：
  矩阵 → 阶梯形 → 简化阶梯形

例1-1：行化简以下矩阵：

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix} = A.$$

①由最左非零列开始，这是个主元列

②在主元列上选一个主元，通过对换变换使之移到顶上：

主元

$$A \overset{r}{\sim} B = \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

主元列

③ 用倍加行变换把主元下面的元素变成 0

$$B \overset{r}{\sim} C = \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

主元

主元列

④ 对"剩下的矩阵"重复上述过程，直到出现阶梯形：

$$C \overset{r}{\sim} D = \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

主元

⑤ 由最右侧主元开始，把每个主元上方的元素变成 0；用倍乘变换把每行的主元变成 1.

$$D \overset{r}{\sim} \begin{bmatrix} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

主元

$$\overset{r}{\sim} \begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

即得到矩阵的简化阶梯形.

- 线性方程组可通过对增广矩阵应用行化简法得到解集的显式表达

例1-2：某个线性方程组已化为简化阶梯形：

$$\left[\begin{array}{ccc:c} 1 & 0 & -5 & 1 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

对应线性方程组：

$$\begin{cases} x_1 \quad -5x_3 = 1 \\ \quad x_2 + x_3 = 4 \\ \qquad\quad 0 = 0 \end{cases}$$

其中 x_1 与 x_2 是基本变量，x_3 是自由变量。

解集的参数表示：

$$\begin{cases} x_1 = 1 + 5x_3 \\ x_2 = 4 - x_3 \end{cases}$$

- 解的存在性与唯一性：

  线性方程组有解，当且仅当增广矩阵的阶梯形最右列不是主元列，即阶梯形中没有形如 $0 \cdots 0 \vdots b \ (b \neq 0)$ 的行。

  若方程组有解：

  i) 当没有自由变量时，有唯一解；

  ii) 当有至少一个自由变量时，有无穷多解。

- 线性方程组的解法总结如下：

  ① 写出增广矩阵

  ② 行化简至阶梯形

  ③ 判断是否有解

  ④ 若有解，继续行化简至简化阶梯形

  ⑤ 写出对应的方程组

  ⑥ 写成用自由变量表示基本变量的形式

1.2 R^n中的向量与矩阵.

- 向量的观点:

  What is a vector?

  Physics: 既有大小,又有方向的量

  CS: n个元素构成的有序数组

  Mathematics: 在定义了满足一定运算律的加法与数乘的集合中的元素.

- 从上节矩阵出发,先定义向量为仅有一列的矩阵.

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

所有包含n个实数的向量构成的集合记为R^n.

- R^n中两个向量相等,当且仅当它们的各对应元素(分量)相等.
- 加法与数乘:

$\forall \vec{u} = (u_1, u_2, \cdots, u_n)^T \in R^n$, $\vec{v} = (v_1, v_2, \cdots, v_n)^T \in R^n$, $\lambda \in R$

定义向量加法 $f: R^n \times R^n \to R^n$, $(\vec{u}, \vec{v}) \mapsto \vec{u} + \vec{v}$

$$\vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2, \cdots, u_n + v_n)^T$$

定义向量数乘 $f: R \times R^n \to R^n$, $(\lambda, \vec{v}) \mapsto \lambda\vec{v}$

$$\lambda\vec{v} = (\lambda v_1, \lambda v_2, \cdots, \lambda v_n)^T$$

记号 ^{T}: 这里表示向量转置

- 零向量: 所有分量均为零的向量,记为0
- 由以上定义,可以易得向量的以下性质.

$\forall \vec{u}, \vec{v}, \vec{w} \in R^n$, $a, b \in R$.

i) 交换性: $\vec{u} + \vec{v} = \vec{v} + \vec{u}$

ii) 结合性: $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$, $(ab)\vec{v} = a(b\vec{v})$

iii) 加法单位元：$\vec{v}+0=\vec{v}$

iv) 加法逆元：$\vec{v}+(-\vec{v})=0$

v) 乘法单位元：$1\cdot\vec{v}=\vec{v}$

vi) 分配性：$a(\vec{u}+\vec{v})=a\vec{u}+a\vec{v}$，$(a+b)\vec{v}=a\vec{v}+b\vec{v}$

- 事实上，任意集合 V 只要定义了满足性质 i)~vi) 的加法与数乘，就构成向量空间，其中的元素称为向量。（详见 2.1）
- 向量的线性组合：

$\forall \vec{v_1},\vec{v_2},\cdots,\vec{v_m}\in R^n$，$\lambda_1,\lambda_2,\cdots,\lambda_m\in R$

则 $\vec{y}=\sum_{i=1}^{m}\lambda_i\vec{v_i}$ 称为向量 $\vec{v_1},\vec{v_2},\ldots,\vec{v_m}$ 以 $\lambda_1,\lambda_2,\ldots,\lambda_m$ 为权的线性组合

- 张成空间：

向量 $\vec{v_1},\vec{v_2},\cdots,\vec{v_m}$ 的所有线性组合构成的集合称为 $\vec{v_1},\vec{v_2},\ldots,\vec{v_m}$ 的张成空间，记为 $span(\vec{v_1},\vec{v_2},\cdots,\vec{v_m})$.

- 引入向量后，1.1 开头的线性方程组可以记为

$$\begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{pmatrix} x_1+\begin{pmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{m,2} \end{pmatrix} x_2+\cdots+\begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{m,n} \end{pmatrix} x_n=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

或 $\sum_{i=1}^{n}\vec{a_i}x_i=\vec{b}$

$\Rightarrow$ $\vec{b}$ 可表示为 $\vec{a_1},\vec{a_2},\ldots,\vec{a_m}$ 的线性组合，当且仅当对应的线性方程组有解.

$\Rightarrow$ $\vec{b}\in span(\vec{a_1},\vec{a_2},\ldots,\vec{a_m})$，当且仅当 ～ .

- 线性方程组可以进一步简记为

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

或 $A\vec{x}=\vec{b}$ （$AX=B$，按矩阵乘法的记号的话）

- 显然以下命题是等价的

A是 $m\times n$ 矩阵

i) $\forall \vec{b}\in R^m$，$A\vec{x}=\vec{b}$ 有解

ii) $\forall \vec{b}\in R^m$，$\vec{b}$ 是A各列的线性组合

iii) A的各列张成 R^m

iv) A在每一行都有主元（详见1.6）

1.3 线性方程组的解集

- 可以写成 $A\vec{x}=0$ 的线性方程组称为齐次线性方程组
- 齐次线性方程组必有解 $\vec{x}=0$，称为平凡解
- $A\vec{x}=0$ 有非平凡解，当且仅当方程的解包含至少一个自由变量.

例1-3：写出 $10x_1-3x_2-2x_3=0$ 的解集.

$$10x_1-3x_2-2x_3=0 \Rightarrow x_1=0.3x_2+0.2x_3$$

$$\Rightarrow \vec{x}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0.3x_2+0.2x_3\\x_2\\x_3\end{pmatrix}=x_2\begin{pmatrix}0.3\\1\\0\end{pmatrix}+x_3\begin{pmatrix}0.2\\0\\1\end{pmatrix}$$

解集也可表示为 $\vec{x}\in span((0.3,1,0)^T,(0.2,0,1)^T)$

其中 $(0.3,1,0)^T$ 与 $(0.2,0,1)^T$ 称为解向量.

- $A\vec{x}=0$ 的解总可以表示为 $span(\vec{v_1},\vec{v_2},\dots,\vec{v_m})$，其中 $\vec{v_1},\vec{v_2},\dots,\vec{v_m}$ 是解向量，若 $A\vec{x}=0$ 仅有平凡解，解为 $span(0)$.
- 非齐次线性方程组 $A\vec{x}=\vec{b}$ （$\vec{b}\neq 0$）

给定 $A\vec{x}=\vec{b}$ 的一个解向量 $\vec{p}$，以及 $A\vec{x}=0$ 的解系 $\vec{v_1},\vec{v_2},\dots,\vec{v_m}$.

$A\vec{x}=\vec{b}$ 的解集可表示为：

$$\vec{x}=\vec{p}+span(\vec{v_1},\vec{v_2},\cdots,\vec{v_m})=\vec{p}+\sum_{i=1}^{m}t_i\vec{v_i} \quad (t_i\in F)$$

1.4 线性无关

- 线性无关的向量组：

  $\forall \vec{v_1}, \vec{v_2}, \dots, \vec{v_m} \in R^n$. 若向量方程 $\sum_{i=1}^{m} x_i\vec{v_i} = 0$ 仅有平凡解，则称向量组 $\vec{v_1}, \vec{v_2}, \dots, \vec{v_m}$ 线性无关.

  若 $\sum_{i=1}^{m} x_i\vec{v_i} = 0$ 有非平凡解，则向量组 $\vec{v_1}, \vec{v_2}, \dots, \vec{v_m}$ 线性相关.

- 矩阵A的各列线性无关，当且仅当 $A\vec{x} = 0$ 仅有平凡解

- 线性无关组的特征：

  i) 一个向量 $\vec{v}$ 构成的向量组线性无关，当且仅当 $\vec{v} \neq 0$

  ii) 两个向量 $\vec{u}, \vec{v}$ 构成的向量组线性无关，当且仅当每个向量都不是另一个向量的标量倍.

  iii) $\vec{v_1}, \vec{v_2}, \dots, \vec{v_m}$ 线性相关，则 $\exists\, \vec{v_j} \in span(\vec{v_1}, \vec{v_2}, \dots, \vec{v_{j-1}})$，且去掉 $\vec{v_j}$ 后的 $\vec{v_1}, \vec{v_2}, \dots, \vec{v_m}$ 仍张成 $span(\vec{v_1}, \vec{v_2}, \dots, \vec{v_m})$ (1-1)

iii) 证：$\vec{v_1}, \vec{v_2}, \dots, \vec{v_m}$ 线性相关 $\Rightarrow \exists$ 不全为零的 $\lambda_1, \lambda_2, \dots, \lambda_m$，使得 $\sum_{i=1}^{m} \lambda_i\vec{v_i} = 0$

令 j 是使 $\lambda_j \neq 0$ 的最大的数.

$\Rightarrow \sum_{i=1}^{j} \lambda_i\vec{v_i} = 0 \Rightarrow \lambda_j\vec{v_j} = -\sum_{i=1}^{j-1}\lambda_i\vec{v_i} \Rightarrow \vec{v_j} = \sum_{i=1}^{j-1}\left(-\frac{\lambda_j}{\lambda_i}\right)\vec{v_i}$

$\Rightarrow \vec{v_j} \in span(\vec{v_1}, \vec{v_2}, \dots, \vec{v_{j-1}})$.

$\forall\, \vec{u} \in span(\vec{v_1}, \vec{v_2}, \dots, \vec{v_j}, \dots, \vec{v_m}) \Rightarrow \vec{u} = \sum_{i=1}^{m} a_i\vec{v_i} = \sum_{i\neq j} a_i\vec{v_i} + a_j\sum_{i=1}^{j-1}\left(-\frac{\lambda_j}{\lambda_i}\right)\vec{v_i}$

$\Rightarrow \vec{u} = \sum_{i\neq j} c_i\vec{v_i} \Rightarrow \vec{u} \in span(\vec{v_1}, \vec{v_2}, \dots, \cancel{\vec{v_j}}, \dots, \vec{v_m})$ QED.

  iii) 更常见的说法为：

  $\vec{v_1}, \vec{v_2}, \dots, \vec{v_m}$ 线性相关 ($m \geq 2$)，当且仅当存在一个向量可以表示为其它向量的线性组合.

  iv) 包含零向量的向量组必然线性相关.

- 线性无关组长度 $\leqslant$ 张成组长度. (1-2)

  设 $\vec{u_1}, \vec{u_2}, \dots, \vec{u_m}$ 在 $span(\vec{v_1}, \vec{v_2}, \dots, \vec{v_n})$ 中线性无关，则 $m \leq n$.

证：$\vec{u_1} \in span(\vec{v_1}, \vec{v_2}, \dots, \vec{v_n}) \Rightarrow \vec{u_1}, \vec{v_1}, \vec{v_2}, \dots, \vec{v_n}$ 是线性相关组. 由定理(1-1)

知，这个组去掉某个$\vec{v}$后，仍能张成span$(\vec{v_1},\vec{v_2},\cdots,\vec{v_n})$，将此时的向量组记为$B_1$；在$B_1$中加入$\vec{u_2}$，去掉某个$\vec{v}$，仍能张成原空间，将向量组记为$B_2$……

重复上述步骤，使得向量组B_m中有所有的$\vec{u}$并仍张成原空间。

这说明$m\leq n$. Q.E.D.

1.5 R^n中的线性映射.

- 映射：集合之间的一种对应关系.

$$f: V\to W,\quad x\mapsto f(x)$$

把集合V中的元素x以规则f对应到W中的元素$f(x)$.

V称为定义域，W称为陪域，$f(x)$称为x的像，所有像$f(x)$构成的集合称为像集或值域.

- 单射：若$f: V\to W$满足$\forall x,y\in V$，$x\neq y\Rightarrow f(x)\neq f(y)$

满射：若$f: V\to W$满足$\forall z\in W$，$\exists x\in V$，$f(x)=z$.

双射 = 单射 + 满射

一些术语说明：
单射 injection = 一对一 one-to-one
满射 surjection = 映上 onto
双射 bijection = 一一对应 one-to-one correspondence

- 线性映射（实向量空间中的）

$T: R^n\to R^m$ 满足

i) $\forall \vec{u},\vec{v}\in R^n$，$T(\vec{u}+\vec{v})=T(\vec{u})+T(\vec{v})$

ii) $\forall \vec{v}\in R^n$，$\lambda\in R$，$T(\lambda\vec{v})=\lambda T(\vec{v})$

则称T为线性映射.

- 令矩阵$A=[\vec{a_1}\ \vec{a_2}\cdots\vec{a_n}]$，则矩阵变换$\vec{x}\mapsto A\vec{x}$是线性映射

1.0

证：设 $\vec{x}=(x_1,x_2,\cdots,x_n)^T\in R^n$，$\vec{u},\vec{v}\in R^n$，$\lambda\in R$

$A\vec{x}=\sum_{i=1}^n \vec{a_i}x_i$

$\Rightarrow A(\vec{u}+\vec{v})=\sum_{i=1}^n \vec{a_i}(u_i+v_i)=\sum_{i=1}^n u_i\vec{a_i}+\sum_{i=1}^n v_i\vec{a_i}=A\vec{u}+A\vec{v}$

$\Rightarrow A(\lambda\vec{v})=\sum_{i=1}^n \vec{a_i}(\lambda v_i)=\lambda\sum_{i=1}^n \vec{a_i}v_i=\lambda A\vec{v}$. 　　2E.D.

• 记号：从集合V到集合W的所有线性映射的集合记为 $\mathcal{L}(V,W)$.

• $\forall T\in\mathcal{L}(R^n,R^m)$，∃唯一的矩阵A，使得 $\forall\vec{x}\in R^n$，$T(\vec{x})=A\vec{x}$.

(1-4)

证：先引入 R^n 中的恒等变换 $I_n: R^n\to R^n$

$\forall\vec{x}\in R^n$，$I_n\vec{x}=\vec{x}$.

I_n 即单位矩阵．$I_n=[\vec{e_1}\ \vec{e_2}\cdots\vec{e_n}]$

第i个元素

$I_n\vec{x}=\vec{x}\Rightarrow\vec{e_i}x_i=x_i\Rightarrow\vec{e_i}=(0,\cdots 1,\cdots 0)$

$$\Rightarrow I_n=\begin{pmatrix}1&0&\cdots&0\\0&1&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&1\end{pmatrix}=diag(1,1,\cdots,1)$$

R^n 中的单位矩阵是 $n\times n$ 矩阵（或称n阶方阵），主对角线（a_{ij} 中 $i=j$ 的元素）全为1，其余元素为0.

在不引起歧义的情况下，I_n 可简记为I.

• Kronecker符号：

$$\delta_{ij}=\begin{cases}1 & i=j\\0 & i\neq j\end{cases}$$

引入该符号后，单位矩阵I可表示为 $\{\delta_{ij}\}$.

(开始证明) $T(\vec{x})=T(I\vec{x})=T(\sum_{i=1}^n x_i\vec{e_i})$

$=\sum_{i=1}^n x_iT(\vec{e_i})=A\vec{x}$

其中 $A=[T(\vec{e_1})\ T(\vec{e_2})\cdots T(\vec{e_n})]$，A是 $m\times n$ 矩阵. 　　2E.D.

• A称为线性映射T的标准矩阵，即标准基下的矩阵.

例1-4：R^2中一些线性映射对应的矩阵.

逆时针旋转φ：$A=\begin{pmatrix}\cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi\end{pmatrix}$

关于x轴对称：$A=\begin{pmatrix}1 & 0 \\ 0 & -1\end{pmatrix}$

关于y轴对称：$A=\begin{pmatrix}-1 & 0 \\ 0 & 1\end{pmatrix}$

关于$x=y$对称：$A=\begin{pmatrix}0 & 1 \\ 1 & 0\end{pmatrix}$

关于原点对称：$A=\begin{pmatrix}-1 & 0 \\ 0 & -1\end{pmatrix}$

投影到x轴：$A=\begin{pmatrix}1 & 0 \\ 0 & 0\end{pmatrix}$

水平剪切：$A=\begin{pmatrix}1 & k \\ 0 & 1\end{pmatrix}$ $(k\in R)$

- 设$T\in\mathcal{L}(R^n,R^m)$，A是T的标准矩阵.
  - i) T是单射，当且仅当A各列线性无关
  - ii) T是满射，当且仅当A各列张成R^m. (1-5)

证：i) 充分性：A各列线性无关 ⇒ $A\vec{x}=0$仅有平凡解

设$\vec{x},\vec{y}\in R^n$使得$A\vec{x}=A\vec{y}$ ⇒ $A(\vec{x}-\vec{y})=0$ ⇒ $\vec{x}-\vec{y}=0$ ⇒ $\vec{x}=\vec{y}$

⇒ A对应的映射T是单射

必要性：T是单射 ⇒ 设$\vec{x}\in R^n$且$T(\vec{x})=T(0)$，则有$\vec{x}=0$

⇒ $A\vec{x}=0$仅有平凡解 ⇒ A各列线性无关.

ii) 充分性：设$A=[\vec{a_1}\ \vec{a_2}\cdots\vec{a_n}]$. A各列张成$R^m$ ⇒

$R^m=\mathrm{span}(\vec{a_1},\vec{a_2},\dots,\vec{a_n})$ ⇒ $\{\vec{y}\in R^m : A\vec{x}=\vec{y}\ |\ \vec{x}\in R^n\}$

$= span(\vec{a_1}, \vec{a_2}, \cdots, \vec{a_n}) = R^m \Rightarrow T$是满射.

必要性：T是满射 $\Rightarrow T$的值域$\{\vec{y} \in R^m : A\vec{x} = \vec{y} \mid \vec{x} \in R^n\} = R^m$.

$\Rightarrow span(\vec{a_1}, \vec{a_2}, \cdots, \vec{a_n}) = R^m \Rightarrow A$的各列张成$R^m$. 之证.

关于双射（可逆，同构）的讨论见2.2节与3.4节.

1.6 主元

- 系数矩阵的主元分析对于线性方程组的解的存在性与唯一性是非常重要的：

  i) $A\vec{x} = \vec{b}$的解（若存在）唯一，当且仅当A的阶梯形在每一列都有主元；

  ii) $A\vec{x} = \vec{b}$的解对$\forall \vec{b}$都存在，当且仅当A的阶梯形在每一行都有主元；

  iii) $A\vec{x} = \vec{b}$的解对$\forall \vec{b}$都存在且唯一，当且仅当A的阶梯形在每一行和每一列都有主元. (1-6)

证：i) 非主元列在方程的解中对应自由变量，方程的解唯一，等价于解中不含自由变量，即等价于系数矩阵的阶梯形不含非主元列.

ii) 设A的阶梯形为A_e.

充分性：若A_e每一行都有主元，则增广矩阵的阶梯形没有形如$0\ 0 \cdots 0 \vdots b\ (b \neq 0)$的行，那么$A\vec{x} = \vec{b}$有解.

必要性：证逆否命题.

设A_e存在非主元行，则A_e的最后一行必全为0.

考虑方程$A_e\vec{x} = \vec{b_e}$，$\vec{b_e} = (0, 0, \cdots, 1)^T$.

显然$A_e\vec{x} = \vec{b_e}$无解，则其对应的原始形式$A\vec{x} = \vec{b}$也无解.

iii) = i) + ii) 之证.

- 定理(1-5)与(1-6)结合即有：

  i) A的各列线性无关，当且仅当A的阶梯形在每一列都有主元.

  ii) A的各列张成R^m，当且仅当A的阶梯形在每一行都有主元. (1-7)

矩阵运算

2.1 矩阵乘法

- 矩阵的记号回顾：

设A是$m\times n$矩阵．$\vec{a}_j$表示A的第j列构成的向量，$a_{i,j}$表示A的第i行第j列的元素．

$$A=[\vec{a}_1\ \vec{a}_2\cdots\vec{a}_n]=\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}=\{a_{ij}\}$$

- 矩阵相等：两个矩阵相等，当且仅当它们行数与列数相同，且对应元素相等．

- 矩阵作为向量的运算：

$m\times n$矩阵本质上可视为mn个数构成的有序数组．也可视为向量．

所有$m\times n$矩阵构成的集合记为$F^{m,n}$．其中F为矩阵定义的数域．

加法与数乘：

$\forall A,B\in F^{m,n}$，$\lambda\in F$．

定义矩阵加法 $f:F^{m,n}\times F^{m,n}\to F^{m,n}$ $(A,B)\mapsto A+B$

$$A+B=\{a_{i,j}\}+\{b_{i,j}\}=\{a_{ij}+b_{ij}\}$$

即对应位置的元素相加．

定义矩阵数乘 $f:F\times F^{m,n}\to F^{m,n}$ $(\lambda,A)\mapsto\lambda A$

$$\lambda A=\lambda\{a_{i,j}\}=\{\lambda a_{i,j}\}$$

即每个元素分别乘以该标量．

容易验证，以上定义的加法与数乘满足与向量相似的运算规律．

※ 实际上这里构造了$F^{m,n}$向量空间，见3.2节讨论．

- 矩阵作为线性映射的运算：

设$T_1\in\mathcal{L}(F^n,F^m)$，$T_2\in\mathcal{L}(F^m,F^p)$．

线性映射的复合 $T_2(T_1(\vec{x})) = (T_2 \circ T_1)(\vec{x})$.

线性映射的复合 $T = T_2 \circ T_1$ 仍是线性映射. $T \in \mathcal{L}(R^n, R^p)$. (2-1)

证：$\forall \vec{u}, \vec{v} \in F^n$, $\lambda \in F$.

$T(\vec{u}+\vec{v}) = T_2(T_1(\vec{u}+\vec{v})) = T_2(T_1(\vec{u}) + T_1(\vec{v}))$

$= T_2(T_1(\vec{u})) + T_2(T_1(\vec{v})) = T(\vec{u}) + T(\vec{v})$.

$T(\lambda\vec{v}) = T_2(T_1(\lambda\vec{v})) = T_2(\lambda T_1(\vec{v})) = \lambda T_2(T_1(\vec{v})) = \lambda T(\vec{v})$ 证毕.

设 T_1 的标准矩阵为 B，T_2 的标准矩阵为 A.

定义矩阵乘积 AB 为 $T = T_2 \circ T_1$ 的标准矩阵，即矩阵乘积对应线性映射的复合.

- 矩阵乘法.

设 $A \in F^{p,m}$，$B \in F^{m,n}$，$\forall \vec{x} \in F^n$.

$$(AB)\vec{x} = A(B\vec{x}) = A\left(\sum_{i=1}^{n} x_i \vec{b_i}\right) = \sum_{i=1}^{n} x_i (A\vec{b_i})$$

$$= [A\vec{b_1}\ A\vec{b_2} \cdots A\vec{b_n}] \cdot \vec{x}$$

$$\Rightarrow AB = [A\vec{b_1}\ A\vec{b_2} \cdots A\vec{b_n}] \quad \alpha$$

其中第 k 列：$A\vec{b_k} = \sum_{j=1}^{m} \vec{a_j} b_{j,k}$.

第 i 行第 k 列：$(A\vec{b_k})_i = \sum_{j=1}^{m} a_{i,j} b_{j,k}$.

若令 $C = AB = \{c_{i,k}\}$，则有

$$c_{i,k} = \sum_{j=1}^{m} a_{i,j} b_{j,k} \quad \beta$$

(或借助 Einstein 求和约定：$c_{i,k} = a_{i,j} b_{j,k}$)

若以下乘积均有定义，则矩阵乘法满足如下性质： (2-2)

i) 结合性：$A(BC) = (AB)C$

ii) 分配性：$A(B+C) = AB + AC$

$(A+B)C = AC + BC$

iii) $\lambda(AB) = (\lambda A)B = A(\lambda B) \quad (\lambda \in F)$

iv) 乘法单位元：$A = AI = IA$

证：ii)~iv) 从定义看是显然的

i) 从线性映射的角度看也是显然的，下面给出代数证明：

设 $A\in F^{q,p}$, $B\in F^{p,m}$, $C\in F^{m,n}$.

这里暂时使用如下记号：$A_{i,j}$ 表示 A 的第 i 行第 j 列的元素.

$$
\begin{aligned}
((AB)C)_{i,l} &= \sum_{k=1}^{m}(AB)_{i,k}C_{k,l} \\
&= \sum_{k=1}^{m}\left(\sum_{j=1}^{p}A_{i,j}B_{j,k}\right)C_{k,l} \\
&= \sum_{j=1}^{p}A_{i,j}\left(\sum_{k=1}^{m}B_{j,k}C_{k,l}\right) \\
&= \sum_{j=1}^{p}A_{i,j}(BC)_{j,l} \\
&= (A(BC))_{i,l}
\end{aligned}
$$

$\Rightarrow (AB)C = A(BC)$. QED.

※ 仅当 A 的列数等于 B 的行数时，AB 才有定义.

※ 矩阵乘法没有交换性，即一般 $AB\neq BA$

（并且 AB 有定义时，BA 不一定有定义）

※ 矩阵乘法没有消去律，即 $AC=BC$ 不能推出 $A=B$.

- 矩阵转置

$\forall A\in F^{m,n}$，A 的转置 $A^T\in F^{n,m}$.

A^T 的列就是 A 的行.

若 $A=\{a_{i,j}\}$，则 $A^T=\{a_{j,i}\}$.

矩阵转置的性质： (2-3)

i) $(A^T)^T = A$

ii) $\forall \lambda_1,\lambda_2,\cdots,\lambda_n\in F$，$\left(\sum_{i=1}^{n}\lambda_iA_i\right)^T=\sum_{i=1}^{n}\lambda_iA_i^T$

iii) $(AB)^T=B^TA^T$.

证：i) 和 ii) 由定义是显然的

iii) 用 $A_{i,j}$ 表示 A 的第 i 行第 j 列元素.

$$((AB)^T)_{i,j}=(AB)_{j,i}=\sum_k A_{j,k}B_{k,i}=\sum_k (B^T)_{i,k}(A^T)_{k,j}$$

1.6.

$= (B^TA^T)_{i,j}$

$\Rightarrow (AB)^T = B^TA^T$ 2ED.

- 若方阵 A 满足 $A = A^T$，则称其为对称矩阵

  $A = -A^T$，则称其为反对称矩阵.

对称矩阵有 $a_{i,j} = a_{j,i}$.

- 算子的幂：

若线性映射的定义域和陪域相同，则称其为线性算子.

从集合 V 映到自身的所有线性算子的集合记为 $\mathcal{L}(V)$.

即 $\mathcal{L}(V) = \mathcal{L}(V,V)$

若 $T \in \mathcal{L}(R^n)$，则 T 的标准矩阵是 n 阶方阵

$\forall n \in N$. $T^n(\vec{x}) = \underbrace{T(T(\cdots T}_{n个T}(x)\cdots))$

对应矩阵的幂 $A^n = \underbrace{AA\cdots A}_{n个A}$ $(A^0 = I_n)$

关于算子的幂的讨论详见5.2

2.2 矩阵的逆

- 设 $A \in F^{m,n}$.

  若 $\exists B \in F^{n,m}$，使 $BA = I_n$，称 A 为左可逆，B 是 A 的左逆

  若 $\exists C \in F^{n,m}$，使 $AC = I_m$，称 A 为右可逆，C 是 A 的右逆

若 A 同时左可逆和右可逆，则称 A 可逆 (1/3)

- 若 A 可逆，则 A 的左逆等于右逆. (2-4)

证：$BAC = B(AC) = BI_m = B$

$BAC = (BA)C = I_nC = C$

$\Rightarrow B = C$ 2ED.

· 可逆矩阵：

重新定义可逆如下： (2/3)

$\forall A \in F^{m,n}$，若∃唯一的 $A^{-1} \in F^{n,m}$，使得 $AA^{-1} = I_m$ 且 $A^{-1}A = I_n$

则称 A 可逆，A^{-1} 称为 A 的逆.

· 可逆矩阵的性质： (2-5)

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

ii) $(AB)^{-1} = B^{-1}A^{-1}$ （若 A、B 可逆）

iii) $(A^T)^{-1} = (A^{-1})^T$

证：i) 是显然的

ii) $(AB)(B^{-1}A^{-1}) = ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I$

$(B^{-1}A^{-1})(AB) = B^{-1}A^{-1}AB = B^{-1}IB = B^{-1}B = I$

$\Rightarrow AB$ 可逆，$(AB)^{-1} = B^{-1}A^{-1}$

iii) $(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$

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

$\Rightarrow A^T$ 可逆，$(A^T)^{-1} = (A^{-1})^T$ QED.

· 矩阵 A 可逆，当且仅当 $A\vec{x} = \vec{b}$ 对 $\forall \vec{b}$ 都有唯一解. (2-6)

证：设 $T \in \mathcal{L}(F^n, F^m)$，$A \in F^{m,n}$ 是 T 的标准矩阵.

充分性：$A\vec{x} = \vec{b}$ 对 $\forall \vec{b}$ 都有唯一解.

$\Rightarrow \forall \vec{b} \in F^m$，∃唯一 $\vec{x} \in F^n$，$T(\vec{x}) = \vec{b}$. 记 $\vec{x} = T'(\vec{b})$.

$T(\lambda_1\vec{x}_1 + \lambda_2\vec{x}_2) = \lambda_1 T(\vec{x}_1) + \lambda_2 T(\vec{x}_2) = \lambda_1\vec{b}_1 + \lambda_2\vec{b}_2$

$\Rightarrow \lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 = T'(\lambda_1\vec{b}_1 + \lambda_2\vec{b}_2)$

$\Rightarrow \lambda_1 T'(\vec{b}_1) + \lambda_2 T'(\vec{b}_2) = T'(\lambda_1\vec{b}_1 + \lambda_2\vec{b}_2)$

$\Rightarrow T' \in \mathcal{L}(F^m, F^n)$. 设 $T'(\vec{b}) = B\vec{b}$. $B \in F^{n,m}$

$\Rightarrow AB\vec{b} = A\vec{x} = \vec{b} = I_m\vec{b} \Rightarrow AB = I_m$

$BA\vec{x} = B\vec{b} = \vec{x} = I_n\vec{x} \Rightarrow BA = I_n$

$\Rightarrow A$可逆，B是A的逆.

必要性：设A可逆，则$A\vec{x}=\vec{b} \Rightarrow A^{-1}A\vec{x}=A^{-1}\vec{b} \Rightarrow \vec{x}=A^{-1}\vec{b}$

$\forall \vec{b}\in F^m$，$\vec{x}=A^{-1}\vec{b}$是方程的唯一解. QED.

※ 若引入基的概念(详见3.2)，该定理可表述为：

矩阵$A\in F^{m,n}$可逆，当且仅当A的各列构成F^m的一组基. (2-7)

· 定理(2-7)可以立刻得到如下结论：

可逆矩阵必为方阵. (2-8)

因为F^m中的一组基长度为m.

· 若方阵A左可逆或右可逆，则A可逆. (2-9)

证：若A左可逆，即$\exists B$使$BA=I$.

对方程$A\vec{x}=0 \Rightarrow BA\vec{x}=B\cdot 0 \Rightarrow \vec{x}=0$.

即$A\vec{x}=0$仅有平凡解 $\Rightarrow A$各列线性无关

$\Rightarrow A$在每一列都有主元.

A是方阵，且每行最多只有一个主元 $\Rightarrow A$在每一行都有主元.

$\Rightarrow A\vec{x}=\vec{b}$的解$\forall\vec{b}$都存在且唯一 $\Rightarrow A$可逆.

若A右可逆，即$\exists C$使$AC=I$.

$\forall \vec{b}$，$\vec{x}=C\vec{b}$. 则$A\vec{x}=AC\vec{b}=I\vec{b}=\vec{b}$.

即$A\vec{x}=\vec{b}$对$\forall\vec{b}$都有解 $\Rightarrow A$在每一行都有主元

A是方阵，且每列最多只有一个主元 $\Rightarrow A$在每一列都有主元

$\Rightarrow A\vec{x}=\vec{b}$的解$\forall\vec{b}$都存在且唯一 $\Rightarrow A$可逆. QED.

※ 注意这个定理仅对方阵成立.

· 有了上述讨论，我们可以最终定义可逆方阵：(3/3)

$\forall A\in F^{n,n}$，若存在唯一的A^{-1}，使得$AA^{-1}=I_n$或$A^{-1}A=I_n$，

则称A可逆，A^{-1}称为A的逆.

· 2阶方阵的逆：

Campus

设 $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$，若 $ad-bc\neq 0$，则 A 可逆，且

$$A^{-1}=\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \qquad (2\text{-}10)$$

这个结论的简单证明见3.3

- 行化简法计算 A^{-1}：

先引入初等矩阵的概念：（列变换等价于右乘初等矩阵）

矩阵 A 进行一次行变换，等价于在 A 上左乘一个初等矩阵 E.

例2-1：3种行变换对应的初等矩阵． (2-11)

i) 对换第 i 行与第 j 行：

$$E_1=\begin{array}{cl}\left(\begin{array}{ccccccccc}
1 &&&&&&&& \\
&\ddots&&&&&&& \\
&&1&&&&&& \\
&&&0&\cdots&1&&& \\
&&&\vdots&1&\vdots&&& \\
&&&\vdots&&1&&& \\
&&&1&\cdots&0&&& \\
&&&&&&1&& \\
&&&&&&&\ddots& \\
&&&&&&&&1
\end{array}\right) & \begin{array}{l} \\ \\ \\ \text{第}i\text{行} \\ \\ \\ \text{第}j\text{行} \\ \\ \\ \\ \end{array}\end{array}$$

第 i 列　第 j 列

（空白处元素全为0，下同）

ii) 第 i 行乘以 λ.

$$E_2=\begin{array}{cl}\left(\begin{array}{ccccccc}
1 &&&&&& \\
&\ddots&&&&& \\
&&1&&&& \\
&&&\lambda&&& \\
&&&&1&& \\
&&&&&\ddots& \\
&&&&&&1
\end{array}\right) & \begin{array}{l} \\ \\ \\ \text{第}i\text{行} \\ \\ \\ \\ \end{array}\end{array}$$

第 i 列

iii) 第 j 行加上第 i 行的 λ 倍.

$$E_3=\begin{array}{cl}\left(\begin{array}{ccccccc}
1 &&&&&& \\
&\ddots&&&&& \\
&&1&&&& \\
&&&\ddots&&& \\
&&\lambda&\cdots&1&& \\
&&&&&\ddots& \\
&&&&&&1
\end{array}\right) & \begin{array}{l} \\ \\ \text{第}i\text{行} \\ \\ \text{第}j\text{行} \\ \\ \\ \end{array}\end{array}$$

第 i 列　第 j 列

※ 行变换可逆，那么初等矩阵都可逆．即 $EE^{-1}=I$．

• n 阶方阵 A 可逆，当且仅当 $A \overset{r}{\sim} I_n$．　(2-12)

证：A 可逆 $\Longleftrightarrow$ A 的阶梯形在每行每列都有主元

$\Longleftrightarrow$ A 的阶梯形的主元构成主对角线 $\Longleftrightarrow$ $A \overset{r}{\sim} I_n$　QED.

• 若 A 可逆，把 A 变成 I 的一系列行变换同时把 I 变成 A^{-1}．　(2-13)

证：A 可逆 $\Rightarrow A \overset{r}{\sim} I$．

∃ 初等矩阵 $E_1, E_2, \cdots, E_p$，使：

$A \overset{r}{\sim} E_1A \overset{r}{\sim} E_2E_1A \overset{r}{\sim} \cdots\cdots \overset{r}{\sim} E_p\cdots E_2E_1A = I$．

$E_1, \cdots, E_p$ 可逆 $\Rightarrow E_p\cdots E_2E_1$ 可逆．

$\Rightarrow A = (E_p\cdots E_2E_1)^{-1}I = (E_p\cdots E_2E_1)^{-1}$

$\Rightarrow A^{-1} = E_p\cdots E_2E_1 = E_p\cdots E_2E_1I$．　QED.

※ 求 A^{-1} 时，写出增广矩阵 $[A \,\vdots\, I]$．

若 $A \overset{r}{\sim} I$，则 $[A \,\vdots\, I] \overset{r}{\sim} [I \,\vdots\, A^{-1}]$．

例 2-2：求矩阵 $A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{pmatrix}$ 的逆．

解：

$$[A \,\vdots\, I] = \left(\begin{array}{ccc:ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array}\right) \overset{r}{\sim} \left(\begin{array}{ccc:ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array}\right)$$

$$\overset{r}{\sim} \left(\begin{array}{ccc:ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{array}\right) \overset{r}{\sim} \left(\begin{array}{ccc:ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{array}\right)$$

$$\overset{r}{\sim} \left(\begin{array}{ccc:ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{3}{2} & -2 & \frac{1}{2} \end{array}\right) \overset{r}{\sim} \left(\begin{array}{ccc:ccc} 1 & 0 & 0 & -\frac{9}{2} & 7 & -\frac{3}{2} \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & \frac{3}{2} & -2 & \frac{1}{2} \end{array}\right)$$

$A \overset{r}{\sim} I$. 所以A可逆. $A^{-1} = \begin{pmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{pmatrix}$

※ 求解方程 $A\vec{x} = \vec{b}$ 时, 行化简 $[A \vdots \vec{b}]$ 的复杂度远小于用上述方式先求 A^{-1} 再算 $\vec{x} = A^{-1}\vec{b}$ (详见2.4对复杂度的讨论)

- 矩阵方程 $AX = B$:

设 $A \in F^{n,n}$ 可逆, $B \in F^{n,p}$. 则方程有唯一解

$$X = A^{-1}B.$$

且 $[A \vdots B] \overset{r}{\sim} [I_n \vdots A^{-1}B]$. (2-14)

证: 设 $E_p \cdots E_2E_1A = I$ $\Rightarrow A^{-1} = E_p \cdots E_2E_1A$

$\Rightarrow E_p \cdots E_2E_1B = A^{-1}B$ $\Rightarrow [A \vdots B] \overset{r}{\sim} [I \vdots A^{-1}B]$ QED.

2.3 矩阵分块

例2-3: 把矩阵A按某种特定方式"切割"分块:

$$A = \left(\begin{array}{ccc:cc:c} 3 & 0 & -1 & 5 & 9 & -2 \\ -5 & 2 & 4 & 0 & -3 & 1 \\ \hdashline -8 & -6 & 3 & 1 & 7 & -4 \end{array}\right)$$

写成 2×3 的分块矩阵:

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \end{pmatrix}$$

它的元素是子矩阵:

$$A_{1,1} = \begin{pmatrix} 3 & 0 & -1 \\ -5 & 2 & 4 \end{pmatrix} \quad A_{1,2} = \begin{pmatrix} 5 & 9 \\ 0 & -3 \end{pmatrix} \quad A_{1,3} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

$$A_{2,1} = \begin{pmatrix} -8 & -6 & 3 \end{pmatrix} \quad A_{2,2} = \begin{pmatrix} 1 & 7 \end{pmatrix} \quad A_{2,3} = \begin{pmatrix} -4 \end{pmatrix}$$

- 矩阵分块.

设$A\in F^{m,n}$. 将A的行分成r段，每段分别含有$m_1, m_2, \cdots, m_r$行；将A的列分成s段，每段分别含有$n_1, n_2, \cdots, n_s$列. 如下：

$$A=\begin{array}{c} \left(\begin{array}{c|c|c|c} & & & \\ \hline & & & \\ \hline \vdots & & & \\ \hline & & & \end{array}\right) \begin{array}{l} \}m_1 \\ \}m_2 \\ \vdots \\ \}m_r \end{array} \\ \underbrace{}_{n_1}\ \underbrace{}_{n_2}\ \cdots\ \underbrace{}_{n_s} \end{array} \qquad \sum_{i=1}^{r} m_i = m \qquad \sum_{j=1}^{s} n_j = n.$$

分块后A可表示为：

$$A=\begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,s} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,s} \\ \vdots & \vdots & \ddots & \vdots \\ A_{r,1} & A_{r,2} & \cdots & A_{r,s} \end{pmatrix}$$

其中$A_{i,j}\in F^{m_i,n_j}$

- 加法&数乘： (2-15)

  设$A, B\in F^{m,n}$，且遵循相同的分块方式.

  则$A+B$的分块也相同，且$A+B=\{A_{i,j}+B_{i,j}\}$.

  $\lambda\in F$，λA的分块同A，且$\lambda A=\{\lambda A_{i,j}\}$.

- 矩阵分块的乘积：

  分块矩阵AB的乘积有定义，仅当A对列的分割与B对行的分割相同.

  设$A=\{A_{ij}\}$，$A_{i,j}\in F^{m_i,n_j}$；$B=\{B_{i,j}\}$，$B_{i,j}\in F^{n_i,k_j}$.

  $AB=C=\{C_{i,j}\}$.

  有$C_{i,j}\in F^{m_i,k_j}$，且$C_{i,j}=\sum_p A_{i,p}B_{p,j}$. (2-16)

  这个结果可以直接从矩阵乘法的定义得到.

- 矩阵乘积的列行展开：

设 $A\in F^{m,n}$，$B\in F^{n,p}$．A的第i列向量记为 A_i，B的第i行向量记为 B^i．

于是 $AB=(A_1A_2\cdots A_n)(B^1B^2\cdots B^n)^T$

$$=\sum_{i=1}^{n}A_iB^i . \quad (2\text{-}17)②$$

上式即矩阵乘积的列行展开式．

※ 目前计算矩阵乘积的方法有 $\alpha.\beta.\gamma$ (P_{14}, P_{14}, P_{23})

• 准对角矩阵．

对角矩阵的乘积是很简单的：

设 $A=diag(a_1,a_2,\cdots,a_n)$，$B=diag(b_1,b_2,\cdots,b_n)$

$\Rightarrow AB=diag(a_1b_1,a_2b_2,\cdots,a_nb_n)$．

（n阶对角矩阵的几何意义是把 F^n 空间的各个方向分别放大某个倍数）．

具有 $diag(A_{1,1},A_{2,2},\cdots A_{s,s})$ 的分块形式的矩阵称为准对角矩阵．其有如下性质：（关于秩的定义见4.5） (2-18)

i) $\forall$ 相同分块的n阶准对角矩阵．

$A=diag(A_1,A_2,\cdots,A_s)$，$B=diag(B_1,B_2,\cdots,B_s)$

$\Rightarrow AB=diag(A_1B_1,A_2B_2,\cdots,A_sB_s)$

ii) $rank\,A=\sum_{i=1}^{s}rank\,A_s$

iii) A可逆 $\Longleftrightarrow A_1,A_2,\cdots,A_s$ 都可逆．

且 $A^{-1}=diag(A_1^{-1},A_2^{-1},\cdots,A_s^{-1})$

证明不给出，见《高等代数简明教程》上册（第二版）P_{145}．

2.4 LU分解

• 三角矩阵：（上三角矩阵/下三角矩阵）

主对角线上方元素全为零的方阵称为下三角矩阵，下方元素全为零的

称为上三角矩阵；

主对角线元素全为1的(上/下)三角矩阵称为单位(上/下)三角矩阵.

- 三角矩阵的乘积也是三角矩阵，且乘积的主对角线元素等于主对角线元素的乘积. (2-19)

证：这里以上三角矩阵为例. 下三角矩阵同理.

设 $A, B\in F^{n,n}$

$$A=\begin{pmatrix} a_1 & * & \cdots & * \\ 0 & a_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix} \quad B=\begin{pmatrix} b_1 & * & \cdots & * \\ 0 & b_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_n \end{pmatrix}$$

记 E_{ij} 为 n 阶方阵，其中第 i 行第 j 列的元素为1，其余元素为0.

$\Rightarrow A=\sum_{1\le i\le j\le n} a_{i,j}E_{i,j}$, $B=\sum_{1\le k\le l\le n} b_{k,l}E_{k,l}$.

$AB=\sum_{1\le i\le j\le n}\sum_{1\le k\le l\le n} a_{i,j}b_{k,l}E_{i,j}E_{k,l}$.

显然有 $E_{i,j}E_{k,l}=E_{i,l}\delta_{j,k}$

$\Rightarrow AB=\sum_{1\le i\le j\le l\le n} a_{i,j}b_{j,l}E_{i,l}$.

其主对角线元素 $i=j=l$ 恰为 a_ib_i.

即：

$$AB=\begin{pmatrix} a_1b_1 & * & \cdots & * \\ 0 & a_2b_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_nb_n \end{pmatrix}$$

QED.

推论：单位三角矩阵的乘积仍为单位三角矩阵.

- LU分解.

设 $A\in F^{m,n}$，且可以仅通过倍加行变换化成阶梯形.

则有 $A=LU$，其中 U 是 A 的阶梯形，L 是一个单位下三角矩阵.

即：

$$A=LU=\begin{pmatrix}1&0&\cdots&0\\ *&1&\cdots&0\\ \vdots&\vdots&\ddots&\vdots\\ *&*&\cdots&1\end{pmatrix}\begin{pmatrix}*&*&\cdots&*\\ 0&*&\cdots&*\\ \vdots&\vdots&\ddots&\vdots\\ 0&0&\cdots&*\end{pmatrix} \quad (2\text{-}20)$$

证：$A \overset{r}{\sim} U \Rightarrow$ ∃初等矩阵 $E_1, E_2, \cdots, E_p$，

使 $E_p \cdots E_2 E_1 A = U$.

$\Rightarrow A = (E_p \cdots E_2 E_1)^{-1} U$.

倍加行变换对应的初等矩阵是单位下三角矩阵.

$\Rightarrow (E_p \cdots E_2 E_1)^{-1} = L$ 是单位下三角矩阵. QED.

- LU分解的算法

①若可能，用倍加变换把A化成阶梯形

②把每次行变换后的主元列依次放入L，其余元素以0填充.

③倍乘变换把L化成单位下三角矩阵.

例2-4：求 $A=\begin{pmatrix}2&4&-1&5&-2\\-4&-5&3&-8&1\\2&-5&-4&1&8\\-6&0&7&-3&1\end{pmatrix}$ 的LU分解.

解：

$$A=\begin{pmatrix}2&4&-1&5&-2\\-4&-5&3&-8&1\\2&-5&-4&1&8\\-6&0&7&-3&1\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}2&4&-1&5&-2\\0&3&1&2&-3\\0&-9&-3&-4&10\\0&12&4&12&-5\end{pmatrix}$$

$$\overset{r}{\sim}\begin{pmatrix}2&4&-1&5&-2\\0&3&1&2&-3\\0&0&0&2&1\\0&0&0&4&7\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}2&4&-1&5&-2\\0&3&1&2&-3\\0&0&0&2&1\\0&0&0&0&5\end{pmatrix}=U$$

把每次行变换的主元列放入 L：

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ -4 & 3 & 0 & 0 \\ 2 & -9 & 2 & 0 \\ -6 & 12 & 4 & 5 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{pmatrix} = L$$

$$\Rightarrow A = LU = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix}$$

※ 若 A 需要通过行对换化成 U，LU 分解时 L 是置换单位下三角矩阵.（单位下三角矩阵行对换后产生的矩阵）.

- LU 分解的应用：

  对于一系列具有相同系数矩阵的线性方程组：

  $A\vec{x} = \vec{b_1}$，$A\vec{x} = \vec{b_2}$，…，$A\vec{x} = \vec{b_p}$.

  为了简化运算，第一个方程通过化简增广矩阵 $[A \mid \vec{b_1}]$ 解出，并同时得到 A 的 LU 分解.

  则后面的方程则变成求解 $L\vec{y_i} = \vec{b_i}$ 与 $U\vec{x_i} = \vec{y_i}$，复杂度降低

- 运算复杂度

  $\forall$ n 阶稠密矩阵 A（大部分元素非零）

  i) 行化简 A 或求 LU 分解的复杂度 $O(\frac{2}{3}n^3)$

  ii) 计算 A^{-1} 复杂度 $O(2n^3)$

  iii) 求乘积 $A^{-1}\vec{b}$ 的复杂度 $O(2n^2)$

  iv) 求解 $L\vec{y} = \vec{b}$ 与 $U\vec{x} = \vec{y}$ 的复杂度 $O(2n^2)$.

行列式

3.1 行列式的性质

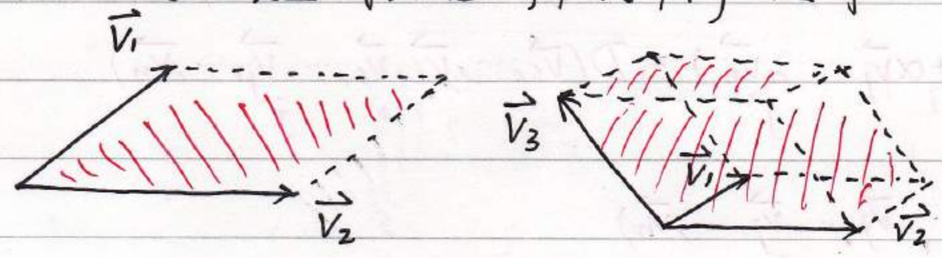
• R^n中的"平行六面体".

考虑R^n中n个向量$\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_n$张成的一定区域:

$$\left\{\vec{v}\in R^n : \vec{v}=\sum_{i=1}^{n} t_i\vec{v}_i \;\middle|\; 0\le t_i\le 1,\ \vec{v}_i\in R^n\right\}$$

当$n=2$时,该区域表示平面上的平行四边形.

当$n=3$时,该区域表示空间中的平行六面体.



$n=2$ $\qquad\qquad$ $n=3$

R^n空间中的"平行六面体"的有向"体积"是$\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_n$的泛函,

记为$D(\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_n)$

从几何的角度考虑,"体积"泛函应当具有如下性质:

① 对每个向量的线性

② 在倍加变换下保持不变.

• 在①的前提下,②等价于反对称性. (3-1)

证: $\forall 1\le i<j\le n$, $D(\vec{v}_1,\cdots,\underset{i}{\vec{v}_i},\cdots,\underset{j}{\vec{v}_j},\cdots,\vec{v}_n)$

$=D(\vec{v}_1,\cdots,\underset{i}{\vec{v}_i},\cdots,\underset{j}{\vec{v}_j-\vec{v}_i},\cdots,\vec{v}_n)$

$=D(\vec{v}_1,\cdots,\underset{i}{\vec{v}_i+(\vec{v}_j-\vec{v}_i)},\cdots,\vec{v}_j-\vec{v}_i,\cdots,\vec{v}_n)$

$=D(\vec{v}_1,\cdots,\underset{i}{\vec{v}_j},\cdots,\underset{j}{(\vec{v}_j-\vec{v}_i)-\vec{v}_j},\cdots,\vec{v}_n)$

$=D(\vec{v}_1,\cdots,\underset{i}{\vec{v}_j},\cdots,\underset{j}{-\vec{v}_i},\cdots,\vec{v}_n)$

$$= -D(\vec{v}_1,\cdots,\underset{i}{\vec{v}_j},\cdots,\underset{j}{\vec{v}_i},\cdots,\vec{v}_n)$$ 倍加变换不变 ⇒ 反对称

另一方面，$D(\vec{v}_1,\cdots,\underset{i}{\vec{v}_i},\cdots,\underset{j}{\vec{v}_j},\cdots,\vec{v}_n) = -D(\vec{v}_1,\cdots,\underset{i}{\vec{v}_j},\cdots,\underset{j}{\vec{v}_i},\cdots,\vec{v}_n)$

$$\Rightarrow D(\vec{v}_1,\cdots,\underset{i}{\vec{v}_i},\cdots,\underset{j}{\vec{v}_i},\cdots,\vec{v}_n) = 0$$

$$D(\vec{v}_1,\cdots,\underset{i}{\vec{v}_i+\alpha\vec{v}_j},\cdots,\underset{j}{\vec{v}_j},\cdots,\vec{v}_n) = D(\vec{v}_1,\cdots,\underset{i}{\vec{v}_i},\cdots,\underset{j}{\vec{v}_j},\cdots,\vec{v}_n)$$

$$+\alpha D(\vec{v}_1,\cdots,\underset{i}{\vec{v}_j},\cdots,\underset{j}{\vec{v}_j},\cdots,\vec{v}_n)$$

$$= D(\vec{v}_1,\cdots,\underset{i}{\vec{v}_i},\cdots,\underset{j}{\vec{v}_j},\cdots,\vec{v}_n)$$ 反对称 ⇒ 倍加变换不变 QED.

- 行列式作为向量组的泛函.

从上述几何直观分析中，抽象出行列式的定义：

$$D: F^n \times F^n \times \cdots \times F^n \to F,\ (\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_n) \mapsto D(\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_n)$$

D满足如下性质：

i) 对每个向量线性：

$$D(\vec{v}_1,\cdots,\alpha\vec{u}_i+\beta\vec{v}_i,\cdots,\vec{v}_n) = \alpha D(\vec{v}_1,\cdots,\vec{u}_i,\cdots,\vec{v}_n) + \beta D(\vec{v}_1,\cdots,\vec{v}_i,\cdots,\vec{v}_n)$$

$\forall\, i \in \{1,2,\cdots,n\}$.

ii) 反对称性：

$$D(\vec{v}_1,\cdots,\underset{i}{\vec{v}_i},\cdots,\underset{j}{\vec{v}_j},\cdots,\vec{v}_n) = -D(\vec{v}_1,\cdots,\underset{i}{\vec{v}_j},\cdots,\underset{j}{\vec{v}_i},\cdots\vec{v}_n)$$

$\forall\, 1 \le i < j \le n$.

iii) 标准化：

$$D(\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n) = 1$$

即行列式是反对称多重线性泛函.

· 行列式作为方阵的泛函.

把 F^n 中的向量 $\vec{v_1}, \vec{v_2}, \cdots, \vec{v_n}$ 作为方阵 A 的列向量:

$A = [\vec{v_1}\ \vec{v_2} \cdots \vec{v_n}]$

定义方阵 A 的行列式 $\det: F^{n,n} \to F$. $A \mapsto \det A$

$\det A = D(\vec{v_1}, \vec{v_2}, \cdots, \vec{v_n})$.

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \text{ 的行列式记为 } \det A = \begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix}$$

· 行列式的性质.

(基本性质)

i) 对每个列向量线性

ii) 交换两个列向量改变行列式的符号.

(导出性质) (3-2)

iii) 若 A 有零列, 则 $\det A = 0$

iv) 若 A 有两列相同, 则 $\det A = 0$.

v) 若 A 不可逆, 则 $\det A = 0$.

证: iii) 在性质 i) 下是显然的; iv) 在性质 ii) 下是显然的.

v). A 不可逆 $\Rightarrow$ A 各列线性相关.

$\Rightarrow$ 若 $A \in F^{n,n} = [\vec{v_1}\ \vec{v_2} \cdots \vec{v_n}]$, 则 $\exists k \in \{1, 2, \cdots, n\}$.

使得 $\vec{v_k} = \sum_{i \neq k}^{n} \lambda_i \vec{v_i}$, $\lambda_i \in F$.

$\Rightarrow \det A = D(\vec{v_1}, \cdots, \underbrace{\sum_{i \neq k}^{n} \lambda_i \vec{v_i}}_{k}, \cdots, \vec{v_n})$

$= \sum_{i \neq k}^{n} \lambda_i D(\vec{v_1}, \cdots, \underbrace{\vec{v_i}}_{k}, \cdots, \vec{v_n})$

求和号内每项都为零（两列相同）$\Rightarrow \det A = 0$　QED.

- 列变换与行列式

若方阵 A 经过一次列变换变为方阵 B，A 与 B 的行列式关系如下：

i) 倍加变换：$\det B = \det A$

ii) 对换变换：$\det B = -\det A$

iii) 倍乘变换：$\det B = \lambda \det A$.

对应的初等矩阵

i) 倍加变换：$\det E = \det I = 1$

ii) 对换变换：$\det E = -\det I = -1$

iii) 倍乘变换：$\det E = \lambda \det I = \lambda$

- 一些特殊矩阵的行列式

对角矩阵的行列式等于主对角线元素乘积.

$$\det[\mathrm{diag}(a_1, a_2, \cdots, a_n)] = a_1 a_2 \cdots a_n. \quad (3\text{-}3)$$

证：$\det[\mathrm{diag}(a_1, a_2, \cdots, a_n)] = D(a_1\vec{e_1}, a_2\vec{e_2}, \cdots, a_n\vec{e_n})$

$= a_1 a_2 \cdots a_n D(\vec{e_1}, \vec{e_2}, \cdots, \vec{e_n}) = a_1 a_2 \cdots a_n$　QED.

三角矩阵的行列式也等于主对角线元素的乘积. (3-4)

$$\begin{vmatrix} a_{1,1} & * & \cdots & * \\ 0 & a_{2,2} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{n,n} \end{vmatrix} = a_{1,1} a_{2,2} \cdots a_{n,n}.$$

证：三角矩阵可以仅通过列倍加变换化成对角矩阵，此过程中行列式不改变.　QED.

- 求方阵行列式的算法.

若方阵 A 不可逆，则 $\det A = 0$.

若方阵 A 可逆，用列变换把 A 化成上三角矩阵 B，B 的对角

线元素为 $b_{1,1}, b_{2,2}, \cdots, b_{n,n}$. 则 $\det B = b_{1,1}b_{2,2}\cdots b_{n,n}$.

设列变换中对换变换的次数为 r，则：

$\det A = (-1)^r \det B = (-1)^r b_{1,1}b_{2,2}\cdots b_{n,n}$.

推论：若 A 可逆，则 $\det A \neq 0$. (3-5)

证：若 A 可逆，上三角矩阵 B 是 A 的阶梯形，则 B 的主对角线元素都是主元，即 $b_{i,i} \neq 0$ $(\forall i \in \{1,2,\cdots,n\}) \Rightarrow \det A \neq 0$. 证毕

- 将 (3-2) 和 (3-5) 结合起来，即有：

  方阵 A 可逆，当且仅当 $\det A \neq 0$. (3-6)

- 引理1：一切可逆矩阵都等价于单位矩阵.

  即 $\forall$ 可逆矩阵 A，$\exists$ 初等矩阵 $E_1, E_2, \cdots, E_p$，

  使得 $A = E_1E_2\cdots E_p$. (3-7)

证：可逆矩阵的阶梯形的主对角线元素都是主元，则简化阶梯形必为单位矩阵.

$\Rightarrow I = E_p' \cdots E_2' E_1' A$.

初等矩阵 E_i' 可逆，令 $(E_i')^{-1} = E_i$

$\Rightarrow A = E_1E_2\cdots E_p I = E_1E_2\cdots E_p$. 证毕.

- 引理2：$\forall$ 可逆矩阵 A & 初等矩阵 E，

  有 $\det(AE) = \det A \cdot \det E$. (3-8)

证：由 P30 结论，显然.

推论：$\det(E_1E_2\cdots E_p) = \det E_1 \cdot \det E_2 \cdot \ldots \cdot \det E_p$.

- 转置不改变行列式.

  $\forall$ 可逆矩阵 A，$\det A = \det A^T$ (3-9)

证：观察3种初等矩阵，易得 $\det E = \det E^T$

设 $A = E_1E_2\cdots E_p$

$\Rightarrow \det A = \det(E_1E_2\cdots E_p) = \det E_1 \cdot \det E_2 \cdot \ldots \cdot \det E_p$

$= \det E_p^T \cdot \det E_{p-1}^T \cdot \ldots \cdot \det E_1^T$

$= \det(E_p^T E_{p-1}^T \cdots E_1^T) = \det(E_1 E_2 \cdots E_p)^T = \det A^T$ QED.

※ 这表明行列式对列变换的所有性质都对行变换适用.

※ 行列式不仅对矩阵的列向量线性,还对行向量线性.

- 方阵乘积的行列式等于行列式的乘积.

$\forall A, B \in F^{n,n}$: $\det(AB) = \det A \cdot \det B$. (3-10)

先给引理:若A或B不可逆且AB有定义,则AB不可逆. (3-11)

证:若A或B非方阵,则AB必不为方阵,即不可逆.

若B为不可逆方阵,假设 $AB = C$ 可逆.

$AB = C \Rightarrow C^{-1}AB = I \Rightarrow B$ 有左逆 $C^{-1}A \Rightarrow B$ 可逆,矛盾.

若A为不可逆方阵,假设 $AB = C$ 可逆 $\Rightarrow B$ 可逆

$\Rightarrow A = CB^{-1}$ 也可逆,矛盾. QED.

再证原结论:

若A或B不可逆,则 $\det A = 0$ 或 $\det B = 0$.

AB也不可逆,$\det(AB) = 0 \Rightarrow \det(AB) = \det A \cdot \det B$.

若A, B均为可逆矩阵,则∃初等矩阵 $E_1, E_2, \cdots, E_p$,

使 $B = E_1 E_2 \cdots E_p$.

$\Rightarrow \det(AB) = \det(AE_1E_2 \cdots E_p)$

$= \det A \cdot \det E_1 \cdot \det E_2 \cdot \ldots \cdot \det E_p$

$= \det A \cdot \det(E_1 E_2 \cdots E_p) = \det A \cdot \det B$. QED.

- 行列式的性质总结如下

i) 行列式对各列向量线性

ii) 行列式对各行向量线性

iii) 交换两个列向量,行列式变号

iv) 交换两个行向量,行列式变号.

v) 若矩阵A有零列，$\det A=0$

vi) 若矩阵A有零行，$\det A=0$

vii) 若矩阵A有两列相同，$\det A=0$

viii) 若矩阵A有两行相同，$\det A=0$

ix) 行列式在倍加变换下保持不变

x) 三角矩阵的行列式等于主对角线元素乘积.

xi) 矩阵A可逆当且仅当 $\det A\neq 0$

xii) $\det A=\det A^T$

xiii) $\det(AB)=\det A\cdot\det B$

xiv) $\forall A\in F^{n,n}$，$\det(\lambda A)=\lambda^n\det A$

xv) 矩阵A满秩当且仅当 $\det A\neq 0$

xvi) 方阵行列式等于各本征值(计重数)乘积.

xvii) 相似方阵的行列式相等.

3.2 行列式的展开

- 从行列式的基本性质推出n阶行列式的完全展开形式：

设 $A\in F^{n,n}$. $A=[\vec{a_1}\ \vec{a_2}\cdots\vec{a_n}]=\{a_{i,j}\}$.

$\vec{a_k}=\sum_{j=1}^{n}a_{j,k}\vec{e_j}$.

于是 $\det A=D(\vec{a_1},\vec{a_2},\cdots,\vec{a_n})$

$=D\left(\sum_{j_1=1}^{n}a_{j_1,1}\vec{e_{j_1}},\sum_{j_2=1}^{n}a_{j_2,2}\vec{e_{j_2}},\cdots,\sum_{j_n=1}^{n}a_{j_n,n}\vec{e_{j_n}}\right)$

$=\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\cdots\sum_{j_n=1}^{n}a_{j_1,1}a_{j_2,2}\cdots a_{j_n,n}D(\vec{e_{j_1}},\vec{e_{j_2}},\cdots,\vec{e_{j_n}})$

考虑 $D(\vec{e_{j_1}},\vec{e_{j_2}},\cdots,\vec{e_{j_n}})$，当任意两角标相同时其值为零.

所有D中(n^n项)，非零D的角标 $j_1,j_2,\cdots,j_n$各异($n!$项).

当$j_1,j_2,\cdots,j_n$构成自然数$1\sim n$的顺序排列$\{1,2,\cdots,n\}$时，

$D=1$，当任意两个j置换时，D变号.

- 将 1~n 从小到大的排列视为标准次序，任一排序 σ 可以由标准次序交换元素得到. 所有 n 个元素的排列的集合记为 $perm(n)$.
某个排列 σ 的对换次数称为排列的逆序数，记为 $r(\sigma)$.
于是 $\forall \sigma \in perm(n)$，$D(\vec{e}_{\sigma_1}, \vec{e}_{\sigma_2}, \cdots \vec{e}_{\sigma_n}) = (-1)^{r(\sigma)} D(\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n)$
$= (-1)^{r(\sigma)}$

$$\Rightarrow \det A = \sum_{\sigma \in perm(n)} (-1)^{r(\sigma)} a_{\sigma_1,1} a_{\sigma_2,2} \cdots a_{\sigma_n,n}. \quad (3\text{-}12)$$

上式即行列式的完全展开形式.

- 余子式展开
余子式：$\forall A \in F^{n,n}$，记 $A_{i,j}$ 为 A 去掉第 i 行第 j 列后剩余的矩阵.
定义余子式 $C_{i,j} = (-1)^{i+j} \det A_{i,j}$
余子式在方阵中的符号如下

$$\begin{pmatrix} + & - & + & \cdots \\ - & + & - & \\ + & - & + & \\ \vdots & & & \ddots \end{pmatrix}$$

- $\forall i \in \{1, 2, \cdots, n\}$，$\det A = \sum_{j=1}^{n} a_{i,j} C_{i,j}$ $\quad (3\text{-}13)$
（行列式按行的余子式展开）.

证：先证 $i=1$，即 $\det A = \sum_{j=1}^{n} a_{1,j} C_{1,j}$
令 $A^{(j)}$ 是 A 中元素 $a_{1,k}$ $(j \neq k)$ 都为零的矩阵.

$$A^{(j)} = \begin{pmatrix} 0 & 0 & \cdots & * & \cdots & 0 \\ * & * & \cdots & * & \cdots & * \\ \vdots & \vdots & & \vdots & & \vdots \\ * & * & \cdots & * & \cdots & * \end{pmatrix}$$

第 j 列

把 $A^{(1)}$ 行化简成下三角矩阵 $U^{(1)}$，注意此过程中元素 $a_{1,1}$ 不改变. 设 $U^{(1)}$ 的主对角线元素为 $a_{1,1}, a_{2,2}, \cdots, a_{n,n}$.

$\Rightarrow \det A^{(1)} = (-1)^r a_{1,1} a'_{2,2} \cdots a'_{n,n}$

其中 r 是行对换次数

而相同的过程把 $A_{1,1}$ 也化成了下三角矩阵 $B_{1,1}$

且 $\det A_{1,1} = (-1)^r a'_{2,2} \cdots a'_{n,n}$

$\Rightarrow \det A^{(1)} = a_{1,1} \det A_{1,1} = a_{1,1} C_{1,1}$

对于 $A^{(2)}$，交换第1列与第2列．把 $A^{(2)}$ 化成类似 $A^{(1)}$ 的形式．

易证 $\det A^{(2)} = -a_{1,2} \det A_{1,2} = a_{1,2} C_{1,2}$

对于 $A^{(3)}$，先交换第3列与第2列，再交换第2列与第1列．

$$\det A^{(3)} = (-1)^2 a_{1,3} \det A_{1,3} = a_{1,3} C_{1,3}$$

以此类推，有 $\det A^{(j)} = (-1)^{j-1} a_{1,j} \det A_{1,j} = a_{1,j} C_{1,j}$．

于是 $\det A = \sum_{j=1}^{n} \det A^{(j)} = \sum_{j=1}^{n} a_{1,j} C_{1,j}$．

当 $i=2$ 时，交换第1行与第2行，化成类似 $i=1$ 的情况．

$$\begin{aligned}\Rightarrow \det A &= -1 \cdot \sum_{j=1}^{n} (-1)^{j-1} a_{2,j} \det A_{2,j} \\ &= -1 \cdot \sum_{j=1}^{n} (-1)^{j+1} a_{2,j} \det A_{2,j} \\ &= \sum_{j=1}^{n} (-1)^{j+2} a_{2,j} \det A_{2,j} \\ &= \sum_{j=1}^{n} a_{2,j} C_{2,j}.\end{aligned}$$

以此类推，最终有 $\forall i \in \{1,2,\cdots,n\}$，$\det A = \sum_{j=1}^{n} a_{i,j} C_{i,j}$ 2ED．

※ 由于转置不改变行列式，那么行列式还可以按列展开：

$$\forall j \in \{1,2,\cdots,n\},\ \det A = \sum_{i=1}^{n} a_{i,j} C_{i,j} \tag{3-14}$$

• 二阶行列式．

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = a|d| - b|c| = ad - bc. \tag{3-15}$$

• 三阶行列式

例3-1：$\forall \vec{v}, \vec{w} \in R^3$，定义 $\vec{v}$ 与 $\vec{w}$ 的叉积：$R^3 \times R^3 \to R^3$，

$(\vec{v}, \vec{w}) \mapsto \vec{v} \times \vec{w}$

$$\vec{v}\times\vec{w}=\begin{vmatrix}\hat{i}&\hat{j}&\hat{k}\\ v_x&v_y&v_z\\ w_x&w_y&w_z\end{vmatrix}$$，其中 $(\hat{i},\hat{j},\hat{k})$ 为 R^3 的标准正交基.

对该行列式按第1行余子式展开：

$$\vec{v}\times\vec{w}=\begin{vmatrix}v_y&v_z\\ w_y&w_z\end{vmatrix}\hat{i}-\begin{vmatrix}v_x&v_z\\ w_x&w_z\end{vmatrix}\hat{j}+\begin{vmatrix}v_x&v_y\\ w_x&w_y\end{vmatrix}\hat{k}$$

$$=(v_yw_z-v_zw_y)\hat{i}+(v_zw_x-v_xw_z)\hat{j}+(v_xw_y-v_yw_x)\hat{k}$$

$$=\begin{pmatrix}v_yw_z-v_zw_y\\ v_zw_x-v_xw_z\\ v_xw_y-v_yw_x\end{pmatrix}$$

例3-2：Vandermonde行列式.

求证：

$$\begin{vmatrix}1&a_0&a_0^2&\cdots&a_0^n\\ 1&a_1&a_1^2&\cdots&a_1^n\\ \vdots&\vdots&\vdots&&\vdots\\ 1&a_n&a_n^2&\cdots&a_n^n\end{vmatrix}=\prod_{0\le i<j\le n}(a_j-a_i)$$

证：对 n 作数学归纳法：

将 a_n 作为变元 x，考虑

$$p(x)=\begin{vmatrix}1&a_0&a_0^2&\cdots&a_0^n\\ 1&a_1&a_1^2&\cdots&a_1^n\\ \vdots&\vdots&\vdots&&\vdots\\ 1&x&x^2&\cdots&x^n\end{vmatrix}$$

对最后一行作余子式展开，则 $p(x)$ 可写作

$p(x)=A_nx^n+A_{n-1}x^{n-1}+\cdots+A_1x+A_0$.

其中 $A_0,A_1,\cdots,A_n$ 是 $a_0,a_1,\cdots,a_{n-1}$ 的函数.

可以看出，$p(x)$ 有零点 $a_0,a_1,\cdots,a_{n-1}$

$\Rightarrow p(x) = A_n(x-a_0)(x-a_1)\cdots(x-a_{n-1})$.

i) 当 $n=1$ 时，Vandermonde 行列式

$$\begin{vmatrix} 1 & a_0 \\ 1 & a_1 \end{vmatrix} = a_1 - a_0 = \prod_{0\le i<j\le 1}(a_j-a_i)$$

原命题成立

ii) 假设 $n=k$ 时命题成立.

即 $$\begin{vmatrix} 1 & a_0 & \cdots & a_0^k \\ 1 & a_1 & \cdots & a_1^k \\ \vdots & \vdots & & \vdots \\ 1 & a_k & \cdots & a_k^k \end{vmatrix} = \prod_{0\le i<j\le k}(a_j-a_i)$$

iii) 当 $n=k+1$ 时.

$p_n(a_{k+1}) = A_{k+1}(a_{k+1}-a_0)(a_{k+1}-a_1)\cdots(a_{k+1}-a_k)$

由余子式展开可以看出 a_{k+1}^{k+1} 前的系数

$$A_{k+1} = (-1)^{(n+2)+(n+2)}\begin{vmatrix} 1 & a_0 & \cdots & a_0^k \\ 1 & a_1 & \cdots & a_1^k \\ \vdots & \vdots & & \vdots \\ 1 & a_k & \cdots & a_k^k \end{vmatrix} = \prod_{0\le i<j\le k}(a_j-a_i)$$

$$\Rightarrow p_n(a_{k+1}) = \prod_{0\le i<j\le k}(a_j-a_i)\cdot\prod_{i=0}^{k}(a_{k+1}-a_i) = \prod_{0\le i<j\le k+1}(a_j-a_i)$$

QED.

· 分块矩阵的行列式

设分块矩阵 $A\in F^{n,n}$，且各分块均为方阵.

$$A = \begin{pmatrix} A_1 & * & \cdots & * \\ 0 & A_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_s \end{pmatrix}$$

有 $\det A = \det A_1 \cdot \det A_2 \cdot \ldots \cdot \det A_s$ (3-16)

证：先证对如下分块矩阵（M_1, M_2为方阵）.

有 $\det M = \det M_1 \cdot \det M_2$

$$M = \begin{pmatrix} M_1 & P \\ 0 & M_2 \end{pmatrix}$$

对M_1的阶数作数学归纳法：

当 $M_1 = c_{1,1}$ 为1阶时，对第1列余子式展开：

$\det M = c_{1,1} \det M_2 = \det M_1 \cdot \det M_2$.

假设 $M_1 \in F^{k,k}$ 时命题成立，当 $M_1 \in F^{k+1,k+1}$ 时，

对第1列余子式展开：

$\det M = \sum_{i=1}^{k+1} (-1)^{i+1} a_{i,1} \det M_{i,1}$.

注意 $M_{i,1} = \begin{pmatrix} (M_1)_{i,1} & * \\ 0 & M_2 \end{pmatrix}$

由归纳假设，$\det M_{i,1} = \det (M_1)_{i,1} \cdot \det M_2$

$\Rightarrow \det M = \underline{\sum_{i=1}^{k+1} (-1)^{i+1} a_{i,1} \det (M_1)_{i,1}} \cdot \det M_2$

$= \det M_1 \cdot \det M_2$.

对A中主对角线子矩阵数s再作归纳法：

$s=1$，$\det A = \det A_1$.

设 $s=k$ 时 $\det A = \prod_{i=1}^{k} \det A_i$

当 $s=k+1$ 时，把以 $A_1, \cdots, A_k$ 为主对角线的区域看作分块 B

则 $\det A = \det B \cdot \det A_{k+1} = \prod_{i=1}^{k+1} \det A_i$. QED.

※ 推论：对n阶准对角矩阵 $\mathrm{diag}(A_1, A_2, \cdots, A_s)$，有：

$\det[\mathrm{diag}(A_1, A_2, \cdots, A_s)] = \det A_1 \cdot \det A_2 \cdot \ldots \cdot \det A_s$. (3-17)

3.3 Cramer 法则

- Cramer 法则.

在求解线性方程组中某个未知量的解时，Cramer 法则可能简化运算

对于方程 $A\vec{x}=\vec{b}$，$A\in F^{n,n}$，$\vec{b}\in F^{n}$.

若 $\det A\neq 0$，则方程有唯一解.

$$x_i=\frac{\det A_i(\vec{b})}{\det A}\quad i\in\{1,2,\cdots,n\} \tag{3-18}$$

其中 $A_i(\vec{b})$ 是系数矩阵 A 的第 i 列被 $\vec{b}$ 替换得到的矩阵.

$$A_i(\vec{b})=[\vec{a}_1\ \vec{a}_2\cdots\underset{i}{\vec{b}}\cdots\vec{a}_n]$$

证：设 $I_i(\vec{x})$ 是单位矩阵 I_n 的第 i 列被 $\vec{x}$ 替换得到的矩阵.

对 $\det I_i(\vec{x})$ 按第 i 列余子式展开：

$\det I_i(\vec{x})=\sum_{k=1}^{n}(-1)^{i+k}x_i\det I_{k,i}$.

当 $k\neq i$ 时，$I_{k,i}$ 只有 $(n-2)$ 个主元，不可逆 $\Rightarrow \det I_{k,i}=0$.

当 $k=i$ 时，$I_{k,i}=I_{n-1}$

$\Rightarrow \det I_{k,i}=\delta_{k,i}$

$\Rightarrow \det I_i(\vec{x})=\sum_{k=1}^{n}(-1)^{i+k}x_i\delta_{k,i}=x_i$

由 $A\cdot I_i(\vec{x})=A_i(\vec{b})$ 得 $\det A\cdot\det I_i(\vec{x})=\det A_i(\vec{b})$.

$\Rightarrow x_i=\det A_i(\vec{b})/\det A$. QED.

- 逆与伴随

※ WARNING："伴随"在线性代数中两个毫不相干的含义，一个表示共轭转置，另一个表示余子式构成的矩阵，注意区分.

伴随矩阵：由 A 的余子式构成的矩阵称为 A 的伴随，记为 A^*（或 adj A）.

有 A^* 的元素 $a^*_{i,j}=C_{j,i}$.

$$\operatorname{adj} A = \begin{pmatrix} C_{1,1} & C_{2,1} & \cdots & C_{n,1} \\ C_{1,2} & C_{2,2} & \cdots & C_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1,n} & C_{2,n} & \cdots & C_{n,n} \end{pmatrix}$$

- 方阵A的$\det A \neq 0$，则A可逆。

  有 $A^{-1} = \frac{1}{\det A} \operatorname{adj} A$. (3-19)

证：设$A^{-1} = [\vec{v}_1\ \vec{v}_2 \cdots \vec{v}_n]$

$AA^{-1} = I_n \Rightarrow A\vec{v}_j = \vec{e}_j$, $j \in \{1, 2, \cdots, n\}$.

由Cramer法则，$\vec{v}_j$的第i个元素：

$v_{i,j} = \frac{\det A_i(\vec{e}_j)}{\det A}$

$\det A_i(\vec{e}_j)$按第i列余子式展开：

$\det A_i(\vec{e}_j) = \sum_{k=1}^{n} e_{k,j} C_{k,i} = \sum_{k=1}^{n} C_{k,i}\delta_{k,j} = C_{j,i}$.

$\Rightarrow v_{i,j} = \frac{1}{\det A} C_{j,i}$

$\Rightarrow A^{-1} = \{v_{i,j}\} = \frac{1}{\det A}\{C_{j,i}\} = \frac{1}{\det A}\operatorname{adj} A$ QED.

- 二阶可逆矩阵的逆。

  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 可逆，当且仅当 $ad-bc \neq 0$.

  $\Rightarrow A^{-1} = \frac{1}{\det A}\begin{pmatrix} C_{1,1} & C_{2,1} \\ C_{1,2} & C_{2,2} \end{pmatrix} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ (3-20)

  这印证了(2-10)的结论。

向量空间

4.1 向量空间与子空间

- 若集合 V 定义了加法与数乘：

  $\forall \vec{u},\vec{v}\in V,\ \vec{u}+\vec{v}\in V$

  $\forall \vec{v}\in V,\ \lambda\in F,\ \lambda\vec{v}\in V$

  且满足如下性质：

  i) 交换性：

  $\forall \vec{u},\vec{v}\in V,\ \vec{u}+\vec{v}=\vec{v}+\vec{u}$

  ii) 结合性：

  $\forall \vec{u},\vec{v},\vec{w}\in V,\ a,b\in F,\ (\vec{u}+\vec{v})+\vec{w}=\vec{u}+(\vec{v}+\vec{w}),\ (ab)\vec{v}=a(b\vec{v})$

  iii) 加法单位元：

  $\exists 0\in V$，使得 $\forall \vec{v}\in V$ 都有 $\vec{v}+0=\vec{v}$

  iv) 加法逆元

  $\forall \vec{v}\in V,\ \exists \vec{w}\in V$ 使 $\vec{v}+\vec{w}=0$

  v) 乘法单位元：

  $\forall \vec{v}\in V,\ 1\cdot\vec{v}=\vec{v}$

  vi) 分配性：

  $\forall \vec{u},\vec{v}\in V,\ a,b\in F,\ a(\vec{u}+\vec{v})=a\vec{u}+a\vec{v},\ (a+b)\vec{v}=a\vec{v}+b\vec{v}$.

  称 V 为向量空间．V 中的元素称为向量

- 其它性质：

  以下性质可由定义易证：

  vii) 加法单位元唯一

  viii) 加法逆元唯一，$\vec{v}$ 的加法逆元记为 $-\vec{v}$.

  ix) $\forall \vec{v}\in V,\ 0\cdot\vec{v}=0$

  x) $\forall \vec{v}\in V,\ (-1)\cdot\vec{v}=-\vec{v}$

- 向量空间的实例

i) F^n

ii) F^S（表示集合 S 到数域 F 的所有函数的集合）

iii) $\mathcal{L}(V, W)$

iv) $F^{m,n}$.

※ F^n 可视为组 $\{1, 2, \cdots, n\}$ 到 F 的函数的集合 (?)

※ 定义了线性映射与矩阵上的加法与数乘，则线性映射集 $\mathcal{L}(V, W)$ 与矩阵集 $F^{m,n}$ 都可视为向量空间．即线性映射或矩阵也可视为向量．

• 线性组合、线性无关 & 张成空间

这些基本概念可以直接从 R^n 迁移到抽象空间 V．

（不啰嗦了）．

• 子空间：

若 V 的子集 U 是向量空间，则称 U 是 V 的子空间．

※ U 上定义的加法与数乘必须与 V 相同．

• V 的子集 U 是 V 的子空间，当且仅当 U 满足如下性质．　(4-2)

i) 加法单位元：$0 \in U$

ii) 加法封闭性：$\forall \vec{u}, \vec{v} \in U \Rightarrow \vec{u} + \vec{v} \in U$．

iii) 数乘封闭性：$\forall \lambda \in F, \vec{v} \in U \Rightarrow \lambda \vec{v} \in U$．

（证明从略）．

※ 加法单位元 $0 \in U$ 等价于 $U \neq \emptyset$

• 张成空间是包含这组向量的最小子空间．　(4-3)

设 $\vec{v_1}, \vec{v_2}, \cdots, \vec{v_m} \in V$，则 $\mathrm{span}(\vec{v_1}, \vec{v_2}, \cdots, \vec{v_m})$ 是包含 $\vec{v_1}, \vec{v_2}, \cdots, \vec{v_m}$ 的最小子空间．

证：首先，$\mathrm{span}(\vec{v_1}, \vec{v_2}, \cdots, \vec{v_m})$ 是 V 的子空间．

i) $0 = 0\vec{v_1} + 0\vec{v_2} + \cdots + 0\vec{v_m} \in \mathrm{span}(\vec{v_1}, \vec{v_2}, \cdots, \vec{v_m})$

ii) 加法封闭：$\sum_{i=1}^{m} a_i\vec{v_i}$, $\sum_{i=1}^{m} b_i\vec{v_i} \in span(\vec{v_1},\vec{v_2},\cdots,\vec{v_m})$.

$\sum_{i=1}^{m} a_i\vec{v_i} + \sum_{i=1}^{m} b_i\vec{v_i} = \sum_{i=1}^{m}(a_i+b_i)\vec{v_i} \in span(\vec{v_1},\vec{v_2},\cdots,\vec{v_m})$

iii) 数乘封闭：$\sum_{i=1}^{m} a_i\vec{v_i} \in span(\vec{v_1},\vec{v_2},\cdots,\vec{v_m})$

$\forall \lambda \in F$, $\lambda\sum_{i=1}^{m} a_i\vec{v_i} = \sum_{i=1}^{m}(\lambda a_i)\vec{v_i} \in span(\vec{v_1},\vec{v_2},\cdots,\vec{v_m})$.

其次，$span(\vec{v_1},\vec{v_2},\cdots,\vec{v_m}) \ni \vec{v_i}$, $\forall i \in \{1,2,\cdots,n\}$.

则 V 包含所有 $\vec{v_i}$ 的子空间必包含 $span(\vec{v_1},\vec{v_2},\cdots,\vec{v_m})$ (由封闭性)，

$span(\vec{v_1},\vec{v_2},\cdots,\vec{v_m})$ 是最小的子空间. QED.

- $\{0\}$ 是 V 的最小子空间，V 本身是 V 的最大子空间.
- 子空间的实例

  i) R^3 的子空间有 $\{0\}$, R^3 , 所有过原点直线 & 所有过原点平面

  ii) R 上全体可微函数的集合是 R^R 的子空间 .
- 有限维向量空间

  若向量空间可由一组有限长度的向量张成，则称其为有限维向量空间
  反之则称为无限维向量空间.

※ 线性代数的研究对象一般限定在有限维向量空间中.

- 有限维向量空间的子空间也是有限维的 . (4-4)

证：设 V 是有限维向量空间，U 是 V 的子空间

i) 若 $U=\{0\}$, 则 U 有限维；若 $U\neq\{0\}$, 取非零向量 $\vec{v_1}\in U$.

ii) 若 $U=span(\vec{v_1})$, 则 U 有限维；若 $U\neq span(\vec{v_1})$, 取非零向量 $\vec{v_2}\in U$
且 $\vec{v_2}\notin span(\vec{v_1})$, 以此类推.

iii) 最后总能找到 $U=span(\vec{v_1},\vec{v_2},\cdots,\vec{v_m})$ 且 $\vec{v_1},\vec{v_2},\cdots,\vec{v_m}$ 线性无关 .

因为由定理 (1-2)，该无关组不长于 V 的任意张成组. 因此 U 有限维. QED

4.2 基与维数

- 基

若V中的一个向量组既线性无关又张成V，则称为V的基.

- V中向量组$\vec{v_1},\vec{v_2},\cdots,\vec{v_n}$是V的基，当且仅当每个$\vec{v}\in V$都能唯一地写成$\vec{v_1},\vec{v_2},\cdots,\vec{v_n}$的线性组合.

即$\forall \vec{v}\in V$，∃唯一的$\lambda_1,\lambda_2,\cdots,\lambda_n\in F$使得$\vec{v}=\sum_{i=1}^{n}\lambda_i\vec{v_i}$. (4-5)

证：充分性 $V=\mathrm{span}(\vec{v_1},\vec{v_2},\cdots,\vec{v_n})$，则∃$\lambda_1,\lambda_2,\cdots,\lambda_n$使$\vec{v}=\sum_{i=1}^{n}\lambda_i\vec{v_i}$.

下证唯一性：又设$\lambda_1',\lambda_2',\cdots,\lambda_n'\in F$使$\vec{v}=\sum_{i=1}^{n}\lambda_i'\vec{v_i}$

$\Rightarrow 0=\sum_{i=1}^{n}\lambda_i\vec{v_i}-\sum_{i=1}^{n}\lambda_i'\vec{v_i}=\sum_{i=1}^{n}(\lambda_i-\lambda_i')\vec{v_i}$.

由$v_1,v_2,\cdots,v_n$线性无关，有$\lambda_i-\lambda_i'=0 \Rightarrow \lambda_i=\lambda_i'$

必要性：若∃唯一的$\lambda_1,\lambda_2,\cdots,\lambda_n$使$\vec{v}=\sum_{i=1}^{n}\lambda_i\vec{v_i}$.

存在性表明$V=\mathrm{span}(\vec{v_1},\vec{v_2},\cdots,\vec{v_n})$.

再令$\vec{v}=0$，由唯一性得$\vec{v_1},\vec{v_2},\cdots,\vec{v_n}$线性无关

即$\vec{v_1},\vec{v_2},\cdots,\vec{v_n}$是V的基. QED.

- 张成组含有基

设$\vec{v_1},\vec{v_2},\cdots,\vec{v_n}$张成V，则可以通过去掉该组中一些向量，使之化成V的基. (4-6)

证：此定理的证明可以参考定理(1-1)：

记$B=\{\vec{v_1},\vec{v_2},\cdots,\vec{v_n}\}$

若$\vec{v_1}=0$，则在B中去掉$\vec{v_1}$.

若$\vec{v_i}\in\mathrm{span}(\vec{v_1},\vec{v_2},\cdots,\vec{v_{i-1}})$，则在B中去掉$\vec{v_i}$，反之则保留.

对$i\in\{2,3,\cdots,n\}$都经如上操作后，B仍张成V；但根据定理(1-1)，B线性无关.则此时B是V的基. QED.

推论：有限维向量空间必有基. (4-7)

(因为有限维向量空间必有张成组，张成组含有基)

- 线性无关组可扩充为基.

设$\vec{v_1},\vec{v_2},\cdots,\vec{v_n}\in V$线性无关，则可以通过添加一些V中的向量

到该组中，使之化成 V 的基. (4-8)

证：设 $\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_m$ 是 V 的基.

考虑 $B = \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n, \vec{u}_1, \vec{u}_2, \cdots, \vec{u}_m\}$，显然 B 张成 V.

对 B 采用定理(4-6)中的操作，把 B 化成 V 的基. 此过程中没有 $\vec{v}_i$ 被去掉（因为线性无关）. 剩下的 B 包含了全部的 $\vec{v}_i$ 和一部分 $\vec{u}_i$（也可能不剩 $\vec{u}_i$），即 $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n$ 可扩充为 V 的基. QED

- 有限维向量空间的任意基长度都相同. (4-9)

证：设 B_1 和 B_2 是有限维向量空间 V 的两组基.

B_1 线性无关而 B_2 张成 V，由定理(1-2)，$length(B_1) \le length(B_2)$

B_2 线性无关而 B_1 张成 V，同理 $length(B_2) \le length(B_1)$.

于是 $length(B_1) = length(B_2)$ QED.

- 维数

定理(4-9)揭示了有限维向量空间中一个非常重要的性质.

有限维向量空间中任意基的长度称为该空间的维数.

记作 $\dim V$.

$\{0\}$ 的维数为零（?）.

- 若 U 是有限维向量空间的子空间，则 $\dim U \le \dim V$. (4-10)

证：U 的基（长度为 $\dim U$）是 V 中的线性无关组，V 的基（长度为 $\dim V$）是 V 中的张成组，由定理(1-2)，$\dim U \le \dim V$. QED.

- 基定理. (4-11)

对有限维向量空间 V：

i) V 中任意长度为 $\dim V$ 的线性无关组都是 V 的基.

ii) V 中任意长度为 $\dim V$ 的张成组都是 V 的基.

证：i) 由(4-8)得线性无关组可扩充成基. 而 $\dim V$ 长的线性无关组的扩充不需要添加任何向量，它自身即成 V 的基.

ii) 由 (4-6) 得张成组可化简成 V 的基. $\dim V$ 长自的张成组的化简不需要去掉任何向量，它自身即成 V 的基. 2&Q.

- 标准基.

F^n 的标准基：

$$\vec{e_1}=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\ \vec{e_2}=\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\ \cdots,\ \vec{e_n}=\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$$

P_n 的标准基：

$\vec{e_0}=1,\ \vec{e_1}=t,\ \vec{e_2}=t^2,\ \cdots,\ \vec{e_n}=t^n$

-※ P_n 是最高为 n 次自的多项式构成的向量空间.

在 P_n 中，微商算子 $\frac{d}{dt}$ 可表示成矩阵：

$\frac{d}{dt}(\vec{e_k})=k\vec{e_{k-1}}\quad (k\neq 0)$

$$\begin{pmatrix}0&1&\cdots&&0\\&0&2&&\vdots\\\vdots&&0&\ddots&\\&&&\ddots&n-1\\0&&\cdots&&0\end{pmatrix}=\frac{d}{dt}$$

4.3 线性映射

- 从 V 到 W 的线性映射是具有如下性质的映射 $T: V\to W$

  i) 加性：$\forall \vec{u},\vec{v}\in V,\ T(\vec{u}+\vec{v})=T(\vec{u})+T(\vec{v})$

  ii) 齐性：$\forall \lambda\in F,\ \vec{v}\in V.\quad T(\lambda\vec{v})=\lambda T(\vec{v})$

记为 $T\in\mathcal{L}(V,W)$ （或 $Hom(V,W)$）

- 线性映射的实例

零映射 $0 \in \mathcal{L}(V,V)$

$\forall \vec{v} \in V,\ 0(\vec{v}) = 0$

※ 注意第一个零是映射，第二个零是V中的加法單元位.

恒等映射 $I \in \mathcal{L}(V,V)$

$\forall \vec{v} \in V,\ I(\vec{v}) = \vec{v}$

多项式微商 $D \in \mathcal{L}(P_n(R), P_n(R))$

$D(p) = p'$

定积分 $T \in \mathcal{L}(P_n(R), R)$

$T(p) = \int_0^1 p(t)\,dt$

- 线性算子是向量空间到自身的线性映射.

记为 $T \in \mathcal{L}(V)$ （或 $End(V)$）

线性泛函是向量空间到数域的线性映射

线性函数是数域到数域的线性映射

- 线性映射 $\mathcal{L}(V,W)$ 作为向量空间

加法 $S, T \in \mathcal{L}(V,W),\ \vec{v} \in V$

$(S+T)(\vec{v}) = S(\vec{v}) + T(\vec{v})$

数乘 $\lambda \in F,\ T \in \mathcal{L}(V,W),\ \vec{v} \in V$

$(\lambda T)(\vec{v}) = \lambda T(\vec{v})$

加法单位元 $0 \in \mathcal{L}(V,W)$

$0(\vec{v}) = 0$

- 线性映射的复合

若 $T \in \mathcal{L}(U,V),\ S \in \mathcal{L}(V,W)$

复合 $S \circ T \in \mathcal{L}(U,W)$：$\forall \vec{u} \in U,\ (S \circ T)(\vec{u}) = S(T(\vec{u}))$

"复合"可以当作"乘积"，性质如下：

i) 结合性：$(T_1 T_2) T_3 = T_1 (T_2 T_3)$

ii) 分配性：$(S_1+S_2)T = S_1T + S_2T$

$S(T_1+T_2) = ST_1 + ST_2$

iii) 单位元：$TI = IT = T$

- 一些有关线性映射的子空间
- 零空间

$\forall T\in \mathcal{L}(V,W)$，$\text{Null}\,T = \{\vec{v}\in V : T(\vec{v}) = 0\}$（或核 $\text{Ker}\,T$）

- 零空间是 V 的子空间 (4-12)

证：i) $T(0) = 0 \Rightarrow 0\in \text{Null}\,T$.

ii) $\forall \vec{u},\vec{v}\in \text{Null}\,T$.

$T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v}) = 0+0 = 0 \Rightarrow \vec{u}+\vec{v}\in \text{Null}\,T$.

加法封闭.

iii) $\forall \vec{v}\in \text{Null}\,T$，$\lambda\in F$

$T(\lambda\vec{v}) = \lambda T(\vec{v}) = \lambda\cdot 0 = 0 \Rightarrow \lambda\vec{v}\in \text{Null}\,T$

数乘封闭.

由定理(4-2)，$\text{Null}\,T$ 是 V 的子空间. QED.

- 单射：若 $T(\vec{u}) = T(\vec{v}) \Rightarrow \vec{u}=\vec{v}$，则 $T\in\mathcal{L}(V,W)$ 是单射.
- T 是单射当且仅当 $\text{Null}\,T = \{0\}$ (4-13)

证：充分性：设 $\text{Null}\,T = \{0\}$

$\forall \vec{u},\vec{v}\in V$ 使 $T(\vec{u}) = T(\vec{v})$.

$\Rightarrow 0 = T(\vec{u}) - T(\vec{v}) = T(\vec{u}-\vec{v}) \Rightarrow \vec{u}-\vec{v}\in \text{Null}\,T$.

$\Rightarrow \vec{u}-\vec{v} = 0 \Rightarrow \vec{u}=\vec{v}$. $\Rightarrow$ T 是单射.

必要性：设 T 是单射

显然 $0\in \text{Null}\,T$. $\forall \vec{v}\in \text{Null}\,T$，$T(\vec{v}) = 0 = T(0)$

$\Rightarrow \vec{v} = 0 \Rightarrow \text{Null}\,T = \{0\}$. QED.

- 值域：$\forall T\in\mathcal{L}(V,W)$，$\text{Ran}\,T = \{T(\vec{v}) : \vec{v}\in V\}$（或像集 $\text{Im}\,T$）

- 值域是W的子空间 (4-14)

证: i) $T(0)=0 \Rightarrow 0 \in Ran\,T$

ii) $\forall \vec{w_1}, \vec{w_2} \in Ran\,T$, $\exists \vec{v_1}, \vec{v_2} \in V$.

$T(\vec{v_1}) = \vec{w_1}$, $T(\vec{v_2}) = \vec{w_2}$

$\Rightarrow T(\vec{v_1}+\vec{v_2}) = T(\vec{v_1}) + T(\vec{v_2}) = \vec{w_1} + \vec{w_2} \in Ran\,T$

加法封闭.

iii) $\forall \vec{w} \in Ran\,T$, $\lambda \in F$. $\exists \vec{v} \in V$.

$T(\vec{v}) = \vec{w}$

$\Rightarrow T(\lambda\vec{v}) = \lambda T(\vec{v}) = \lambda\vec{w} \in Ran\,T$.

数乘封闭

由定理(4-2), $Ran\,T$是W的子空间. QED.

- 满射: 若$Ran\,T = W$, 则$T \in \mathcal{L}(V,W)$是满射.
- 线性映射基本定理 (4-15)

设V是有限维向量空间. $T \in \mathcal{L}(V,W)$. 则$Ran\,T$是有限维的, 且

$\dim V = \dim Null\,T + \dim Ran\,T$

证: 设$\vec{u_1}, \vec{u_2}, \cdots, \vec{u_m}$是$Null\,T$的基. 于是$\dim Null\,T = m$.

把它扩充为V的基$\vec{u_1}, \vec{u_2}, \cdots, \vec{u_m}, \vec{v_1}, \vec{v_2}, \cdots, \vec{v_n}$. 于是$\dim V = m+n$

接下去只要证$T(\vec{v_1}), T(\vec{v_2}), \cdots, T(\vec{v_n})$是$Ran\,T$的基.

$\forall \vec{v} \in V$, $\exists a_1, \cdots, a_m, b_1, \cdots, b_n \in F$.

$\vec{v} = \sum_{i=1}^{m} a_i\vec{u_i} + \sum_{j=1}^{n} b_j\vec{v_j}$

$\Rightarrow T(\vec{v}) = \sum_{i=1}^{m} a_i T(\vec{u_i}) + \sum_{j=1}^{n} b_i T(\vec{v_j}) = \sum_{j=1}^{n} b_i T(\vec{v_j})$.

上式表明$T(\vec{v_1}), \cdots, T(\vec{v_n})$张成$Ran\,T$, $Ran\,T$是有限维的

设$c_1, c_2, \cdots, c_n \in F$使得$\sum_{j=1}^{n} c_i T(\vec{v_i}) = 0$

$\Rightarrow T(\sum_{j=1}^{n} c_j\vec{v_j}) = 0 \Rightarrow \sum_{j=1}^{n} c_j\vec{v_j} \in Null\,T$.

由于$\vec{u_1}, \vec{u_2}, \cdots, \vec{u_m}$张成$Null\,T$, $\exists d_1, d_2, \cdots, d_m \in F$

使得 $\sum_{j=1}^{n} c_j \vec{v_j} = \sum_{i=1}^{m} d_i \vec{u_i}$

$\vec{v_1}, \cdots, \vec{v_n}, \vec{u_1}, \cdots, \vec{u_m}$ 线性无关，则 c_j 和 d_i 全等于 0.

于是 $T(\vec{v_1}), T(\vec{v_2}), \cdots, T(\vec{v_n})$ 线性无关

$\Rightarrow T(\vec{v_1}), T(\vec{v_2}), \cdots, T(\vec{v_n})$ 是 $\mathrm{Ran}\, T$ 的基. $\dim \mathrm{Ran}\, T = n$.

$\Rightarrow \dim V = \dim \mathrm{Null}\, T + \dim \mathrm{Ran}\, T$ QED.

推论1：到低维空间的线性映射不是单射. (4-16)

证：设 $T \in \mathcal{L}(V, W)$ 且 $\dim V > \dim W$

$\dim \mathrm{Null}\, T = \dim V - \dim \mathrm{Ran}\, T$

$\geq \dim V - \dim W > 0$. QED.

推论2：到高维空间的线性映射不是满射. (4-17)

证：设 $T \in \mathcal{L}(V, W)$ 且 $\dim V < \dim W$

$\dim \mathrm{Ran}\, T = \dim V - \dim \mathrm{Null}\, T$

$\leq \dim V$

$< \dim W$ QED.

4.4 同构

- 线性映射可逆的概念类似于矩阵可逆

线性映射 $T \in \mathcal{L}(V, W)$ 可逆，若存在 $S \in \mathcal{L}(W, V)$ 使得 $ST = I_V$ 且 $TS = I_W$.

易证 S 唯一，称为 T 的逆，记作 $S = T^{-1}$.

- 可逆性等价于双射. (4-18)

证：i) 设 $T \in \mathcal{L}(V, W)$ 可逆

$\forall \vec{u}, \vec{v} \in V$ 使得 $T(\vec{u}) = T(\vec{v})$

$\Rightarrow \vec{u} = T^{-1}(T(\vec{v})) = (T^{-1}T)(\vec{v}) = \vec{v}$ $\Rightarrow T$ 是单射.

$\forall \vec{w} \in W$, $\vec{w} = (TT^{-1})(\vec{w}) = T(T^{-1}(\vec{w})) \in \mathrm{Ran}\, T$.

$\Rightarrow W \subset \mathrm{Ran}\,T$. 又 $\mathrm{Ran}\,T \subset W$ $\Rightarrow \mathrm{Ran}\,T = W$ $\Rightarrow$ T是满射

于是T是双射

ii) 若$T \in \mathcal{L}(V,W)$是双射 .

T是双射 ,令$S(\vec{w}) \in V$是唯一使$T(S(\vec{w})) = \vec{w}$的元素 .

$\Rightarrow TS = I_W$.

$\forall \vec{v} \in V,\ T((ST)(\vec{v})) = (TS)T(\vec{v}) = I_W T(\vec{v}) = T(\vec{v})$

$\Rightarrow (ST)(\vec{v}) = \vec{v}$ $\Rightarrow ST = I_V$.

最后还要证S是线性的 :

$\forall \vec{w_1}, \vec{w_2} \in W,\ T(S(\vec{w_1}) + S(\vec{w_2})) = (TS)\vec{w_1} + (TS)\vec{w_2} = \vec{w_1} + \vec{w_2}$

$\Rightarrow S(\vec{w_1}) + S(\vec{w_2}) = S(\vec{w_1} + \vec{w_2})$.

$\forall \vec{w} \in W,\ \lambda \in F.\ T(\lambda S(\vec{w})) = \lambda (TS)(\vec{w}) = \lambda \vec{w}$

$\Rightarrow \lambda S(\vec{w}) = S(\lambda \vec{w})$.

即S是线性的 .因此T可逆 , $T^{-1} = S$.

综上 ,线性映射可逆 ,当且仅当它是双射 . QED .

• 同构 .

对线性映射 ,同构就是可逆 .

对两个向量空间 ,若它们之间存在同构映射 ,则向量空间同构

※ "同构"反映了两个向量空间本质上相同 .

• 有限维向量空间(F上的)同构当且仅当维数相等 . (4-19)

证 : i) 若V与W同构 . 设$T \in \mathcal{L}(V,W)$可逆 .

$\Rightarrow T$是双射 $\Rightarrow \mathrm{Null}\,T = \{0\}$. $\mathrm{Ran}\,T = W$

$\Rightarrow \dim \mathrm{Null}\,T = 0$. $\dim \mathrm{Ran}\,T = \dim W$

$\Rightarrow \dim V = \dim \mathrm{Null}\,T + \dim \mathrm{Ran}\,T = \dim W$

ii) 若$\dim V = \dim W$

设$\vec{v_1}, \vec{v_2}, \cdots, \vec{v_n}$是$V$的基 , $\vec{w_1}, \vec{w_2}, \cdots, \vec{w_n}$是$W$的基 .

令 $T \in \mathcal{L}(V, W)$. $T(\sum_{i=1}^{n} c_i \vec{v_i}) = \sum_{i=1}^{n} c_i \vec{w_i}$.

$\vec{w_1}, \vec{w_2}, \cdots, \vec{w_n}$张成 $W \Rightarrow Ran\, T = W \Rightarrow T$是满射.

$\vec{w_1}, \vec{w_2}, \cdots, \vec{w_n}$线性无关 $\Rightarrow Null\, T = \{0\} \Rightarrow T$是单射

$\Rightarrow T$是双射. 可逆. 同构 $\Rightarrow V$与W同构. QED.

※ 以上定理表明任意F上的n维向量空间都同构于F^n.

- 设 $\dim V = n$, $\dim W = m$. 有

  $\mathcal{L}(V, W)$与$F^{m,n}$同构. (4-20)

证: 设$\vec{v_1}, \vec{v_2}, \cdots, \vec{v_n}$是$V$的基, $\vec{w_1}, \vec{w_2}, \cdots, \vec{w_m}$是$W$的基.

$\forall j \in \{1, 2, \cdots, n\}$, 设 $T(\vec{v_j}) = \sum_{i=1}^{m} a_{i,j} \vec{w_i}$.

$$记\ M(T) = \{a_{i,j}\} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}$$

于是$M(T)$构成了从$\mathcal{L}(V, W)$到$F^{m,n}$的线性映射.

若$M(T) = 0 \Rightarrow \forall i,j,\ a_{i,j} = 0 \Rightarrow T(\vec{v_j}) = 0$.

$\vec{v_j}$是V的基 $\Rightarrow \forall \vec{v} \in V$, $T(\vec{v}) = 0 \Rightarrow T = 0 \Rightarrow M(T)$是单射

$M(T)$是满射很显然. 则$M(T)$是双射. 可逆. 同构

$\Rightarrow \mathcal{L}(V, W)$与$F^{m,n}$同构. QED

※ 该定理揭示了线性映射空间与矩阵空间的深刻关系. 事实上, 矩阵就是线性映射的表示, 只要向量空间取定基(引入坐标系). 更详细的讨论见4.5

- $\dim F^{m,n} = mn$ (4-21)

证: 考虑矩阵$E_{i,j}$(第i行第j列为1, 其余位置为0), $E_{1,1}, \cdots, E_{m,n}$恰好构成$F^{m,n}$的基. 这个基的长度为mn. $\dim F^{m,n} = mn$. QED

- $\dim \mathcal{L}(V, W) = \dim V \cdot \dim W$ (4-22)

证：$\mathcal{L}(V,W)$ 与 $F^{m,n}$ 同构.

$\dim\mathcal{L}(V,W) = \dim F^{m,n} = mn = \dim V\cdot\dim W$. 2ED

- 线性算子 $\mathcal{L}(V)$
- 算子单射等价于满射. (4-23)

证：i) 若 $T\in\mathcal{L}(V)$ 是单射. $\mathrm{Null}\,T=\{0\}$

$\dim\mathrm{Ran}\,T = \dim V - \dim\mathrm{Null}\,T = \dim V$

$\Rightarrow T$ 是满射.

ii) 若 $T\in\mathcal{L}(V)$ 是满射. $\mathrm{Ran}\,T = V$

$\dim\mathrm{Null}\,T = \dim V - \dim\mathrm{Ran}\,T = 0$

$\Rightarrow \mathrm{Null}\,T=\{0\} \Rightarrow T$ 是单射. 2ED

※ 该定理仅对有限维向量空间成立！

对无限维向量空间，反例容易构造：

$P(\mathbb{R})$ 上的映射 $f(p(x)) = x^2p(x)$ 是单射而不是满射；

F^∞ 上的映射 $(x_1,x_2,x_3,\cdots)\mapsto(x_2,x_3,x_4,\cdots)$ 是满射而不是单射.

4.5 坐标系与基变换

- 设 $\mathcal{B}=\{\vec{b}_1,\vec{b}_2,\cdots,\vec{b}_n\}$ 是 V 的基. $\forall\vec{v}\in V$, $\exists$ 唯一的 $x_1,x_2,\cdots,x_n$.

$\vec{v}=\sum_{i=1}^{n}x_i\vec{b}_i$

$x_1,x_2,\cdots,x_n$ 称为 $\vec{v}$ 在 $\mathcal{B}$ 中的坐标. $\vec{v}$ 相对 $\mathcal{B}$ 的坐标向量

$[\vec{v}]_{\mathcal{B}} = (x_1,x_2,\cdots,x_n)^T\in F^n$

- V 到 F^n 的同构映射 $\vec{v}\mapsto[\vec{v}]_{\mathcal{B}}$ 称为坐标映射.

它也把基 $\vec{b}_1,\vec{b}_2,\cdots,\vec{b}_n$ 映到 F^n 的标准基.

- 设 $T\in\mathcal{L}(V,W)$. $\mathcal{B}=\{\vec{b}_1,\vec{b}_2,\cdots,\vec{b}_n\}$ 是 V 的基，$\mathcal{A}=\{\vec{a}_1,\vec{a}_2,\cdots,\vec{a}_m\}$ 是 W 的基.

由第1章结论，F^n 到 F^m 的线性映射可以用 $F^{m,n}$ 中的矩阵表示.

$\mathcal{L}(V,W)$ 与 $F^{m,n}$ 同构，即 T 也可以用 $m\times n$ 矩阵 A 表示.

$[\vec{w}]_{\mathcal{A}} = [M]_{\mathcal{A}\mathcal{B}}[\vec{v}]_{\mathcal{B}}$.

其中 $[M]_{\mathcal{A}\mathcal{B}}$ 是 T 在基 $\mathcal{A}$ 与 $\mathcal{B}$ 下的矩阵.

- 矩阵乘法表示线性映射的复合.

$T_1 \in \mathcal{L}(V,W)$, $T_2 \in \mathcal{L}(U,V)$. $\mathcal{C}$ 是 U 的基, $\mathcal{B}$ 是 V 的基, $\mathcal{A}$ 是 W 的基. 线性映射的复合 $(T_1T_2)(\vec{u})$ 表示为

$[\vec{w}]_{\mathcal{A}} = [M]_{\mathcal{A}\mathcal{B}}[\vec{v}]_{\mathcal{B}} = [M_1]_{\mathcal{A}\mathcal{B}}[M_2]_{\mathcal{B}\mathcal{C}}[\vec{u}]_{\mathcal{C}}$

即 $[M]_{\mathcal{A}\mathcal{C}} = [M_1]_{\mathcal{A}\mathcal{B}}[M_2]_{\mathcal{B}\mathcal{C}}$.

- 基变换

设 $\mathcal{A} = \{\vec{a}_1, \vec{a}_2, \cdots, \vec{a}_n\}$ 和 $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \cdots, \vec{b}_n\}$ 是 V 的两组基.

考虑恒等映射 I_V 在基 $\mathcal{A}$ 与 $\mathcal{B}$ 下的矩阵 $[I]_{\mathcal{A}\mathcal{B}}$

$[\vec{v}]_{\mathcal{A}} = [I]_{\mathcal{A}\mathcal{B}}[\vec{v}]_{\mathcal{B}}$.

$[I]_{\mathcal{A}\mathcal{B}}$ 称为由 $\mathcal{B}$ 到 $\mathcal{A}$ 的过渡矩阵. 它把 $\forall \vec{v}\in V$ 在基 $\mathcal{B}$ 下的坐标变换到基 $\mathcal{A}$ 下的坐标. 过渡矩阵可逆, 因为列向量线性无关.

- 求解过渡矩阵.

$[I]_{\mathcal{A}\mathcal{B}}$ 把 $[\vec{b}_1]_{\mathcal{B}}, \cdots, [\vec{b}_n]_{\mathcal{B}}$ (即标准基) 映到 $[\vec{b}_1]_{\mathcal{A}}, \cdots, [\vec{b}_n]_{\mathcal{A}}$.

也即 $[I]_{\mathcal{A}\mathcal{B}} = [[\vec{b}_1]_{\mathcal{A}}\ [\vec{b}_2]_{\mathcal{A}} \cdots [\vec{b}_n]_{\mathcal{A}}]$

特殊地, 对于 F^n 中的两组基.

$\mathcal{A} = \{\vec{a}_1, \vec{a}_2, \cdots, \vec{a}_n\}$, $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \cdots, \vec{b}_n\}$

标准基 $\mathcal{S} = \{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\}$

$[I]_{\mathcal{S}\mathcal{B}} = [[\vec{b}_1]_{\mathcal{S}}\ [\vec{b}_2]_{\mathcal{S}} \cdots [\vec{b}_n]_{\mathcal{S}}] = [\vec{b}_1\ \vec{b}_2 \cdots \vec{b}_n] = B$

$[I]_{\mathcal{B}\mathcal{S}} = [I]_{\mathcal{S}\mathcal{B}}^{-1} = B^{-1}$.

增广矩阵 $[A \,\vdots\, B] = [[I]_{\mathcal{S}\mathcal{A}} \,\vdots\, [I]_{\mathcal{S}\mathcal{B}}]$

$\to [[I]_{\mathcal{A}\mathcal{S}}[I]_{\mathcal{S}\mathcal{A}} \,\vdots\, [I]_{\mathcal{A}\mathcal{S}}[I]_{\mathcal{S}\mathcal{B}}] = [I \,\vdots\, [I]_{\mathcal{A}\mathcal{B}}]$

即把 A 化成 I 的初等行变换把 B 换成过渡矩阵 $[I]_{\mathcal{A}\mathcal{B}}$.

- 相似矩阵
- 设$T\in \mathcal{L}(V,W)$. $\mathcal{B}$与$\mathcal{B}'$是V的两组基，$\mathcal{A}$与$\mathcal{A}'$是W的两组基
T在$\mathcal{A}$与$\mathcal{B}$下的矩阵$[M]_{\mathcal{A}\mathcal{B}}$和在$\mathcal{A}'$与$\mathcal{B}'$下的矩阵$[M]_{\mathcal{A}'\mathcal{B}'}$遵循类似矩阵乘的关系：

$$[M]_{\mathcal{A}'\mathcal{B}'}=[I]_{\mathcal{A}'\mathcal{A}}[M]_{\mathcal{A}\mathcal{B}}[I]_{\mathcal{B}\mathcal{B}'}$$

（注意角标关系）

- 设$T\in\mathcal{L}(V)$，$\mathcal{A}$与$\mathcal{B}$是V的两组基.
同样地，$[M]_{\mathcal{A}\mathcal{A}}=[I]_{\mathcal{A}\mathcal{B}}[M]_{\mathcal{B}\mathcal{B}}[I]_{\mathcal{B}\mathcal{A}}$
$[M]_{\mathcal{A}\mathcal{A}}$，$[M]_{\mathcal{B}\mathcal{B}}$是$T$在基$\mathcal{A}$、$\mathcal{B}$下的矩阵，分别记为$A$、$B$，
$[I]_{\mathcal{A}\mathcal{B}}$是过渡矩阵，记为$P$.
$\Rightarrow A=PBP^{-1}$
- 若对于$A.B\in F^{n,n}$，$\exists P\in F^{n,n}$可逆，使得$A=PBP^{-1}$，则称A与B是相似矩阵. $B=P^{-1}AP$
- 相似矩阵反映了同一线性算子在不同基下的表现形式，它们之间用过渡矩阵连接.
- 若$A=PBP^{-1}$，则$\forall n\in N$，$A^n=PB^nP^{-1}$ (4-24)

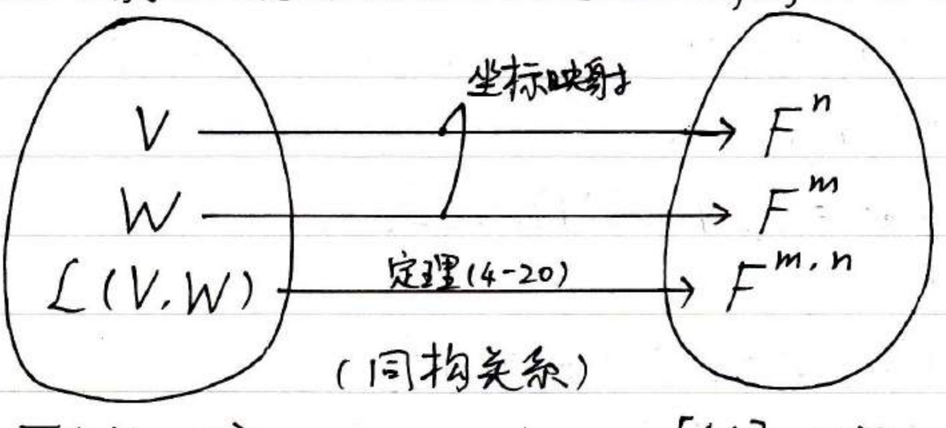
证：对n作归纳
$n=0$时. $A^0=PB^0P^{-1}=I$
假设$n=k$时，$A^n=PB^nP^{-1}$
当$n=k+1$时，$A^n=A^k\cdot A=(PB^kP^{-1})(PBP^{-1})=P(B^k\cdot B)P^{-1}$
$=PB^{k+1}P^{-1}$. 归纳假设成立. 证毕.

- 一点个人的注解.

关于线性映射与矩阵的关系的问题一直困扰着初学者. 在一开始的从R^n到R^m的线性映射，矩阵几乎完全等同于线性映射. 但对抽象向量空间从V到W的线性映射却不然.

对于矩阵，$f(\vec{x}) = A\vec{x}$，$A \in F^{m,n}$. 这里 $f \in \mathcal{L}(F^n, F^m)$.

线性映射取定基，本质上就是建立同构：



$T(\vec{v}) = \vec{w}$ $\qquad\qquad$ $[M]_{AB}[\vec{v}]_B = [\vec{w}]_A$

4.6 矩阵的基本子空间

- 对于 $T \in \mathcal{L}(V, W)$ 在某个基下的矩阵 A.

  零空间 $\mathrm{Null}\,A = \{\vec{v} \in F^n \mid A\vec{v} = 0\} \subset F^n$

  值域 $\mathrm{Ran}\,A = \{\vec{w} \in F^m \mid \vec{w} = A\vec{v},\ \vec{v} \in F^n\} \subset F^m$

  矩阵的值域 $\mathrm{Ran}\,A$ 又称矩阵的列空间 $\mathrm{Col}\,A$. 因为 A 的值域就是 A 的各列向量张成的空间.

- 对 A 的转置 A^T

  A^T 的值域 $\mathrm{Ran}\,A^T$ 称为 A 的行空间 $\mathrm{Row}\,A$

  A^T 的零空间 $\mathrm{Null}\,A^T$ 称为 A 的左零空间（没符号）

- 零空间 $\mathrm{Null}\,A$，左零空间 $\mathrm{Null}\,A^T$，列空间 $\mathrm{Ran}\,A$，行空间 $\mathrm{Ran}\,A^T$ 称为矩阵 A 的4个基本子空间（不是矩阵空间的子空间）

- 求子空间的基

- 零空间的基

  把 $A\vec{x} = 0$ 的解表示成参数向量形式，即求出了 $\mathrm{Null}\,A$ 的基.

例4-1

$$A=\begin{pmatrix}-3&6&-1&1&-7\\1&-2&2&3&-1\\2&-4&5&8&-4\end{pmatrix}$$

$A\vec{x}=0 \Rightarrow \vec{x}=x_2(2,1,0,0,0)^T+x_4(1,0,-2,1,0)^T+x_5(-3,0,2,0,1)^T$

即 Null A 的一组基为 $(2,1,0,0,0)^T$, $(1,0,-2,1,0)^T$, $(-3,0,2,0,1)^T$.

- 列空间的基

矩阵A的主元列（即对应阶梯形中有主元的列）构成 Ran A 的基. (4-25)

证: A的简化阶梯形 A_{re} 的主元列构成 Ran A_{re} 的基.

$A_{re}=E_1E_2\cdots E_pA$

行变换不改变线性相关，则A的主元列是线性无关；

对A的任一列向量 $\vec{v}$，A化成 A_{re} 时 $\vec{v}$ 化成 $E_1E_2\cdots E_p\vec{v}$.

表示为 A_{re} 主元列的线性组合，

$E_1E_2\cdots E_p\vec{v}=\sum_i \lambda_i E_1E_2\cdots E_p\vec{v}_i$

行变换可逆 $\Rightarrow \vec{v}=\sum_i\lambda_i\vec{v}_i$. 即A的主元列张成 Ran A.

$\Rightarrow$ A的主元列构成 Ran A 的基. QED

- 行空间的基

两个行等价的矩阵行空间相同.

阶梯形矩阵的非零行构成行空间的基. (4-26)

证: i) $\forall A \overset{r}{\sim} B$，B的行可以表示为A的行的线性组合，于是 $\mathrm{Ran}B^T\subset \mathrm{Ran}A^T$

反之有 $\mathrm{Ran}A^T\subset\mathrm{Ran}B^T \Rightarrow \mathrm{Ran}A^T=\mathrm{Ran}B^T$.

ii) 应用定理(4-6)中的步骤（从下至上），剩下的线性无关组即B的非零行，又B的非零行张成 $\mathrm{Ran}B^T$，于是B的非零行构成 $\mathrm{Ran}B^T$ 的基.

例4-2

$$A=\begin{pmatrix}1&1&2&2&1\\2&2&1&1&1\\3&3&3&3&2\end{pmatrix}$$

$$A \sim \begin{pmatrix} 1 & 1 & 2 & 2 & 1 \\ 0 & 0 & -3 & -3 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

主元列　　主元列

$Ran A$ 的一组基是 $(1,2,3)^T, (2,1,3)^T$

$Ran A^T$ 的一组基是 $(1,1,2,2,1)^T, (0,0,-3,-3,-1)^T$.

4.7 秩

- 行秩与列秩.

  列秩是矩阵列空间的维数 $\dim Ran A$

  行秩是矩阵行空间的维数 $\dim Ran A^T$

- 行变换不改变行秩；列变换不改变列秩. (4-27)

  由定理(4-26)以及定义立得.

- 秩定理：行秩等于列秩. (4-28)

  $$\dim Ran A = \dim Ran A^T$$

证：由定理(4-25)，$\dim Ran A$ 是 A 中主元列的个数，也即 A 的阶梯形 U 中主元列的个数；由定理(4-26)，$\dim Ran A^T$ 是 U 中非零行的个数. 此两者都等于 B 中主元个数.

因此 $\dim Ran A = \dim Ran A^T$　　Q.E.D.

- 秩的定义

  秩是矩阵列空间的维数.

  $$rank A = \dim Ran A$$

- 矩阵乘与秩

  $\forall A \in F^{p,m}, B \in F^{m,n}$

  $$rank A + rank B - m \le rank(AB) \le \min\{rank A, rank B\} \quad (4\text{-}29)$$

证: i) 设 $C=AB=[A\vec{b}_1\ A\vec{b}_2 \cdots A\vec{b}_n]$

C的列 $\vec{c}_k=\sum_j \vec{a}_j b_{j,k}$ 可以表示为A的列的线性组合.

即 $\mathrm{Ran}\,C \subset \mathrm{Ran}\,A \Rightarrow \dim \mathrm{Ran}\,C \le \dim \mathrm{Ran}\,A$

$\Rightarrow \mathrm{rank}\,C \le \mathrm{rank}\,A$.

取转置: $C^T=B^TA^T$. 同理有 $\mathrm{rank}\,C^T \le \mathrm{rank}\,B^T$.

行秩等于列秩, $\mathrm{rank}\,C \le \mathrm{rank}\,B$

$\Rightarrow \mathrm{rank}\,C \le \min\{\mathrm{rank}\,A,\ \mathrm{rank}\,B\}$.

ii) 设 $T\in \mathcal{L}(\mathrm{Ran}B,\ \mathrm{Ran}AB)$. $T(B\vec{x})=AB\vec{x}$.

由线性映射基本定理,

$$\dim \mathrm{Ran}\,B=\dim \mathrm{Null}\,T+\dim \mathrm{Ran}\,T.$$

$\mathrm{Ran}\,T \subset \mathrm{Ran}\,AB \Rightarrow \dim \mathrm{Ran}\,T \le \mathrm{rank}(AB)$.

$\mathrm{Null}\,T=\{\vec{x}\in \mathrm{Ran}\,B \mid A\vec{x}=0\}$

$\mathrm{Null}\,A=\{\vec{x}\in F^m \mid A\vec{x}=0\}$.

$\Rightarrow \mathrm{Ran}\,B \subset F^m \Rightarrow \mathrm{Null}\,T \subset \mathrm{Null}\,A \Rightarrow \dim \mathrm{Null}\,T \le \dim \mathrm{Null}\,A$.

$\Rightarrow \mathrm{rank}\,B \le \dim \mathrm{Null}\,A+\mathrm{rank}(AB)$.

再由线性映射基本定理, $\dim \mathrm{Null}\,A=m-\dim \mathrm{Ran}\,A$

代入上式得 $\mathrm{rank}\,B \le m-\mathrm{rank}\,A+\mathrm{rank}(AB)$

即 $\mathrm{ran}A+\mathrm{rank}\,B-m \le \mathrm{rank}(AB)$. Q.E.D.

- 满秩

$\forall A\in F^{n,n}$. 若 $\mathrm{rank}\,A=n$, 则称A满秩.

- 矩阵可逆当且仅当满秩. (4-30)

证一: 设 $A=[\vec{a}_1\ \vec{a}_2 \cdots \vec{a}_n]$, $\mathrm{Ran}\,A=\mathrm{span}(\vec{a}_1,\vec{a}_2,\cdots,\vec{a}_n)$.

A可逆 $\Leftrightarrow \vec{a}_1,\vec{a}_2,\cdots,\vec{a}_n$ 是 $\mathrm{Ran}\,A$ 的基 $\Leftrightarrow \dim \mathrm{Ran}\,A=n$

$\Leftrightarrow \mathrm{rank}\,A=n \Leftrightarrow$ A满秩. Q.E.D.

证二: 由线性映射基本定理, $n=\dim \mathrm{Null}\,A+\mathrm{rank}\,A$

A是方阵：

A可逆 $\Leftrightarrow$ 线性映射 $\vec{x} \mapsto A\vec{x}$ 是单射 $\Leftrightarrow \dim Null A = 0$

$\Leftrightarrow rank A = n \Leftrightarrow$ A满秩. Q.E.D.

- 子式与秩

$\forall A \in F^{m,n}$, 在其中选出k行k列，构成k阶子矩阵，该子矩阵的行列式称为k阶子式.

- 非零矩阵的秩等于该矩阵最大非零子式的阶数. (4-31)

证：若 $k > rank A$，由于 $Ran A$ 中基的长度为k，则A中任意k列必线性相关，即k阶子矩阵各列线性相关，k阶子式为0；

若 $k = rank A$，取出A的所有主元行与主元列构成k阶子矩阵. 此子矩阵可逆（定理(1-6)+定理(2-6)），k阶子式非零. Q.E.D.

4.8 子空间的和

- 交、并、和

设 U_1、U_2 是向量空间V的子集.

交：$U_1 \cap U_2 = \{\vec{v} \in V \mid \vec{v} \in U_1 \ \&\& \ \vec{v} \in U_2\}$

并：$U_1 \cup U_2 = \{\vec{v} \in V \mid \vec{v} \in U_1 \ || \ \vec{v} \in U_2\}$

和：$U_1 + U_2 = \{\vec{u_1} + \vec{u_2} \mid \vec{u_1} \in U_1, \vec{u_2} \in U_2\}$

- 子空间的交与和仍是子空间，但子空间的并未必是子空间.

（证明从略）

- 子空间的和是包含这些子空间的最小子空间. (4-32)

证：显然 $\forall U_i \in \sum_j U_j$，反之V中任何包含 $U_1, \dots, U_m$ 的子空间都包含 $\sum_j U_j$（子空间包含其中元素的所有有限和）. 即 $\sum_j U_j$ 是V中包含 $U_1, \dots, U_m$ 的最小子空间. Q.E.D.

- 子空间的直和.

若 $U_1+U_2+\cdots+U_m$ 中每个元素 $\vec{u}$ 都可以唯一表示成 $\vec{u}_1+\vec{u}_2+\cdots+\vec{u}_m$，其中 $\vec{u}_1\in U_1, \vec{u}_2\in U_2, \cdots, \vec{u}_m\in U_m$，那么 $U_1+U_2+\cdots+U_m$ 称为直和，记作 $U_1\oplus U_2\oplus\cdots\oplus U_m=\bigoplus_{i=1}^{m}U_i$.

- 设 $U_1, U_2, \cdots, U_m$ 是 V 的子空间. $U_1+U_2+\cdots+U_m$ 是直和，当且仅当 $0=\sum_{i=1}^{m}\vec{u}_i\ (\vec{u}_i\in U_i)\Rightarrow$ 所有 $\vec{u}_i=0$. (4-33)

证：充分条件显然，只证必要条件：

若 $0=\sum_{i=1}^{m}\vec{u}_i\Rightarrow$ 所有 $\vec{u}_i=0$.

设 $\vec{v}\in U_1+U_2+\cdots+U_m$, $\vec{v}=\sum_{i=1}^{m}\vec{u}_i\ (\vec{u}_i\in U_i)$

再令 $\vec{v}=\sum_{i=1}^{m}\vec{v}_i\ (\vec{v}_i\in U_i)$

两式相减得 $0=\sum_{i=1}^{m}(\vec{u}_i-\vec{v}_i)\Rightarrow\vec{u}_i=\vec{v}_i$.

即 $\vec{v}$ 的线性表示唯一，$U_1+U_2+\cdots+U_m$ 是直和. Q.E.D.

- 设 U 与 W 是 V 的子空间. $U+W$ 是直和，当且仅当 $U\cap W=\{0\}$. (4-34)

证：充分性：若 $U\cap W=\{0\}$.

设 $0=\vec{u}+\vec{w}\ (\vec{u}\in U, \vec{w}\in W)$. $\vec{w}=-\vec{u}\in U\Rightarrow\vec{w}\in U\cap W$

$\Rightarrow\vec{w}=0$. 同理 $\vec{u}=0$. 即 $U+W$ 是直和

必要性：若 $U+W$ 是直和

设 $\vec{v}\in U\cap W$. 有 $0=\vec{v}+(-\vec{v})$, $\vec{v}\in U, -\vec{v}\in W$.

$U+W$ 是直和 $\Rightarrow\vec{v}=0\Rightarrow U\cap W=\{0\}$. Q.E.D.

- $\forall$ 有限维向量空间 V 中的子空间 U，$\exists$ V 的子空间 W，使得 $V=U\oplus W$. (4-35)

证：U 的基 $\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_m$ 是 V 中的线性无关组，可以扩充为 V 的基：

$\vec{u}_1, \cdots, \vec{u}_m, \vec{w}_1, \cdots, \vec{w}_n$.

令 $W=\mathrm{span}(\vec{w}_1, \vec{w}_2, \cdots, \vec{w}_n)$

$\forall\vec{v}\in V$, $\exists a_1, \cdots, a_m, b_1, \cdots, b_n\in F$，使得

$\vec{v}=\sum_{i=1}^{m}a_i\vec{u}_i+\sum_{j=1}^{n}b_j\vec{w}_j=\vec{u}+\vec{w}\ (\vec{u}\in U, \vec{w}\in W)$.

$\Rightarrow \vec{v} \in U + W \Rightarrow V = U + W$

再证直和：$\forall \vec{v} \in U \cap W$，$\exists a_1, \cdots, a_m, b_1, \cdots, b_n \in F$，使得

$\vec{v} = \sum_{i=1}^{m} a_i \vec{u_i} = \sum_{j=1}^{n} b_j \vec{w_j}$

$\Rightarrow \sum_{i=1}^{m} a_i \vec{u_i} + \sum_{j=1}^{n} (-b_j) \vec{w_j} = 0$.

$\vec{u_1}, \cdots, \vec{u_m}, \vec{w_1}, \cdots, \vec{w_n}$ 线性无关 $\Rightarrow a_i = 0, b_j = 0. \ \forall i, j$

$\Rightarrow \vec{v} = 0 \Rightarrow U \cap W = \{0\} \Rightarrow V = U \oplus W$ Q.E.D.

推论：$V = U \oplus W$，当且仅当 U 的基与 W 的基可以并成 V 的基。(4-36)

- 和空间的维数.

设 U_1 和 U_2 是有限维向量空间 V 的子空间，则：

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2) \tag{4-37}$$

证：设 $U_1 \cap U_2$ 的基为 $\vec{u_1}, \vec{u_2}, \cdots, \vec{u_m} \Rightarrow \dim(U_1 \cap U_2) = m$

将其扩充为 U_1 的基 $\vec{u_1}, \cdots, \vec{u_m}, \vec{v_1}, \cdots, \vec{v_r} \Rightarrow \dim U_1 = m + r$

和 U_2 的基 $\vec{u_1}, \cdots, \vec{u_m}, \vec{w_1}, \cdots, \vec{w_s} \Rightarrow \dim U_2 = m + s$.

只要证 $\dim(U_1 + U_2) = m + r + s$，即 $\vec{u_1}, \cdots, \vec{u_m}, \vec{v_1}, \cdots, \vec{v_r}, \vec{w_1}, \cdots, \vec{w_s}$ 是 $U_1 + U_2$ 的基.

$U_1, U_2 \subset \operatorname{span}(\vec{u_1}, \cdots, \vec{u_m}, \vec{w_1}, \cdots, \vec{w_s}, \vec{v_1}, \cdots, \vec{v_r})$

$\Rightarrow U_1 + U_2 \subset \operatorname{span}(\vec{u_1}, \cdots, \vec{u_m}, \vec{w_1}, \cdots, \vec{w_s}, \vec{v_1}, \cdots, \vec{v_r})$

再证线性无关：

$\exists a_1, \cdots, a_m, b_1, \cdots, b_r, c_1, \cdots, c_s \in F$，使得 $\sum_{i=1}^{m} a_i \vec{u_i} + \sum_{j=1}^{r} b_j \vec{v_j} + \sum_{k=1}^{s} c_k \vec{w_k} = 0$

$\Rightarrow \sum_{k=1}^{s} c_k \vec{w_k} = -\sum_{i=1}^{m} a_i \vec{u_i} - \sum_{j=1}^{r} b_j \vec{v_j} \in U_1$

又 $\vec{w_1}, \vec{w_2}, \cdots, \vec{w_s} \in U_2 \Rightarrow \sum_{k=1}^{s} c_k \vec{w_k} \in U_1 \cap U_2$

$\Rightarrow \exists d_1, d_2, \cdots, d_m \in F$，使得 $\sum_{k=1}^{s} c_k \vec{w_k} = \sum_{i=1}^{m} d_i \vec{u_i}$.

$\vec{u_1}, \cdots, \vec{u_m}, \vec{w_1}, \cdots, \vec{w_s}$ 线性无关 $\Rightarrow c_k = 0, d_i = 0 \quad \forall i, k$.

$\Rightarrow \sum_{i=1}^{m} a_i \vec{u_i} + \sum_{j=1}^{r} b_j \vec{v_j} = 0$.

$\vec{u_1}, \cdots, \vec{u_m}, \vec{v_1}, \cdots, \vec{v_r}$ 线性无关 $\Rightarrow a_i = 0, b_j = 0 \quad \forall i, j$

$\Rightarrow \vec{u_1},\cdots,\vec{u_m},\vec{v_1},\cdots,\vec{v_r},\vec{w_1},\cdots,\vec{w_s}$ 线性无关.

最终有 $\dim(U_1+U_2)=\dim U_1+\dim U_2-\dim(U_1\cap U_2)$ Q.E.D

推论: $\dim(U_1\oplus U_2)=\dim U_1+\dim U_2$. (4-38)

4.9 积空间与商空间

· 向量空间的积.

设 $V_1,V_2,\cdots,V_m$ 是同一数域上的向量空间, 积

$V_1\times V_2\times\cdots\times V_m=\{(\vec{v_1},\vec{v_2},\cdots,\vec{v_m})\mid \vec{v_1}\in V_1,\vec{v_2}\in V_2,\cdots,\vec{v_m}\in V_m\}$

规定 $V_1\times V_2\times\cdots\times V_m$ 上的

加法: $(\vec{u_1},\vec{u_2},\cdots,\vec{u_m})+(\vec{v_1},\vec{v_2},\cdots,\vec{v_m})=(\vec{u_1}+\vec{v_1},\vec{u_2}+\vec{v_2},\cdots,\vec{u_m}+\vec{v_m})$

数乘: $\lambda(\vec{v_1},\vec{v_2},\cdots,\vec{v_m})=(\lambda\vec{v_1},\lambda\vec{v_2},\cdots,\lambda\vec{v_m})$.

则易证 $V_1\times V_2\times\cdots\times V_m$ 构成向量空间.

※ $R^2\times R^3$ 与 R^5 同构 (但不相等)

$R^2\times R^3$ 的元素是 $((x_1,x_2),(x_3,x_4,x_5))$

R^5 的元素是 (x_1,x_2,x_3,x_4,x_5)

· 积的维数等于维数的和.

设 $V_1,V_2,\cdots,V_m$ 是有限维向量空间, 则 $V_1\times V_2\times\cdots\times V_m$ 也是有限维的. 且

$\dim(V_1\times V_2\times\cdots\times V_m)=\sum_{i=1}^{m}\dim V_i$ (4-39)

证: 选取每个 V_i 的一组基.

对每个 V_i 的一个基向量 $\vec{v_{ij}}$, 考虑 $\vec{u_k}\in V_1\times V_2\times\cdots V_m$,

$\vec{u_k}=(0,0,\cdots,\underset{i}{\vec{v_{ij}}},\cdots,0)$. 这样的 $\vec{u_k}$ 有 $\sum_{i=1}^{m}\dim V_i$ 个, 且构成

$V_1\times V_2\times\cdots\times V_m$ 的基. 即 $\dim(V_1\times V_2\times\cdots\times V_m)=\sum_{i=1}^{m}\dim V_i$. Q.E.D.

· 积与直和

设 $U_1,U_2,\cdots,U_m$ 是 V 的子空间. 定义 $\Gamma\in\mathcal{L}(U_1\times\cdots\times U_m,U_1+\cdots+U_m)$,

$\Gamma(\vec{u_1},\cdots,\vec{u_m})=\vec{u_1}+\cdots+\vec{u_m}$.

则Γ是单射，当且仅当$U_1+U_2+\cdots+U_m$是直和． (4-40)

证：Γ是单射$\Leftrightarrow$ $\text{Null}\,\Gamma=\{0\}$ $\Leftrightarrow(0=\vec{u_1}+\cdots+\vec{u_m}\Rightarrow\vec{u_i}=0,\ \forall i)$.

$\Leftrightarrow U_1+U_2+\cdots+U_m$是直和． Q.E.D.

- $U_1+U_2+\cdots+U_m$是直和，当且仅当$\dim(U_1+U_2+\cdots+U_m)=\sum_{i=1}^{m}\dim U_i$ (4-41)

证：可以由(4-38)直接得到，也可以通过Γ来辅助证明：

由映射Γ的定义知它是满射．

于是Γ是单射$\Leftrightarrow$ Γ是同构$\Leftrightarrow U_1\times\cdots\times U_m$与$U_1+\cdots+U_m$同构．

$\Leftrightarrow\dim(U_1\times\cdots\times U_m)=\dim(U_1+\cdots+U_m)$

$\Leftrightarrow\dim(U_1+\cdots+U_m)=\sum_{i=1}^{m}\dim U_i$ 再结合(4-40). Q.E.D.

- 商空间

把高维向量空间分解成低维子空间的直和是研究向量空间结构的重要手段之一；而另一方法是引入商空间，把向量空间按照某种性质"分类"，每类用一个向量来代表，从而达到降维的目的．

- 仿射子集 （或同余类）

在V中定义如下关于子空间U的关系：

若$\vec{v}-\vec{w}\in U$，则$\vec{v}\sim\vec{w}$

易证其满足反身性、对称性、传递性，则该关系是一个等价关系．

所有与$\vec{v}$等价的向量可表示为：

$$\vec{v}+U=\{\vec{v}+\vec{u}\mid\vec{v}\in V,\ \vec{u}\in U\}\subset V$$

$\vec{v}+U$称为V中的一个仿射子集，它平行于U.

(某些国内教材称之为"模U的同余类")

- 平行于U的仿射子集或相等或不相交．

设U是V的子空间，$\forall\vec{v},\vec{w}\in V$，以下表述等价： (4-42)

i) $\vec{v}-\vec{w}\in U$

ii) $\vec{v}+U=\vec{w}+U$

iii) $(\vec{v}+U)\cap(\vec{w}+U)\neq\phi$

证：使用轮转证法：

i)⇒ii)：若 $\vec{v}-\vec{w}\in U$，$\forall\vec{u}\in U$，

$\vec{v}+\vec{u}=\vec{w}+[(\vec{v}-\vec{w})+\vec{u}]\subset\vec{w}+U$

$\Rightarrow\vec{v}+U\subset\vec{w}+U$．同理有 $\vec{w}+U\subset\vec{v}+U$

$\Rightarrow\vec{v}+U=\vec{w}+U$

ii)⇒iii)：U 是子空间 $\Rightarrow U\neq\phi\Rightarrow\vec{v}+U\neq\phi$

$\Rightarrow(\vec{v}+U)\cap(\vec{w}+U)=\vec{v}+U\neq\phi$

iii)⇒i)：若 $(\vec{v}+U)\cap(\vec{w}+U)\neq\phi$．则 $\exists\vec{u}_1,\vec{u}_2\in U$．

使得 $\vec{v}+\vec{u}_1=\vec{w}+\vec{u}_2$

$\Rightarrow\vec{v}-\vec{w}=\vec{u}_2-\vec{u}_1\in U$． Q.E.D.

※ V中的任意向量都属于某个仿射子集，而不同的仿射子间不相交，于是V就可以看成相互分离的仿射子集的并集．

- 商空间 V/U．

设U是V的子空间．商空间 V/U 是所有平行于U的仿射子集的集合

$V/U=\{\vec{v}+U\mid\vec{v}\in V\}$

接下来在 V/U 上引入加法和数乘，使之成为向量空间：

加法：$\forall\vec{v},\vec{w}\in V$．$(\vec{v}+U)+(\vec{w}+U)=(\vec{v}+\vec{w})+U$．

这里证明该加法是良定义的（这不显然，因为同一仿射子集有不同的表示方法）．

即 $\forall\vec{v},\vec{w}\in V$．假设 $\vec{v}',\vec{w}'\in V$ 使得 $\vec{v}+U=\vec{v}'+U$ 和 $\vec{w}+U=\vec{w}'+U$．须证明 $(\vec{v}+\vec{w})+U=(\vec{v}'+\vec{w}')+U$：

$\vec{v}+U=\vec{v}'+U$，$\vec{w}+U=\vec{w}'+U$．

⇒由定理(4-42)，$\vec{v}-\vec{v}'=U$，$\vec{w}-\vec{w}'+U$．

$\Rightarrow U$加法封闭，$(\vec{v}-\vec{v'})+(\vec{w}-\vec{w'})\in U$

$\Rightarrow$由定理(4-42)，$(\vec{v}+\vec{w})+U=(\vec{v'}+\vec{w'})+U$. Q.E.D.

数乘：$\forall \vec{v}\in V, \lambda\in F$, $\lambda(\vec{v}+U)=\lambda\vec{v}+U$

这里证明该数乘是良定义的(理由同前)

即$\forall \vec{v}\in V$，假设$\vec{v'}\in V$使得$\vec{v}+U=\vec{v'}+U$.

须证明 $\lambda\vec{v}+U=\lambda\vec{v'}+U$：

$\vec{v}+U=\vec{v'}+U \Rightarrow \vec{v}-\vec{v'}\in U \Rightarrow \lambda(\vec{v}-\vec{v'})=\lambda\vec{v}-\lambda\vec{v'}\in U$

$\Rightarrow \lambda\vec{v}+U=\lambda\vec{v'}+U$. Q.E.D.

有了加法与数乘，易证V/U是向量空间.

注意V/U上的加法单位元是U. $\vec{v}+U$的加法逆元是$-\vec{v}+U$.

例4-3：设R^2上的子空间$U=\{(x,2x)\mid x\in R\}$.

则U是斜率$k=2$且过原点的直线；

设$\vec{v}=(\pi,e)$，则$\vec{v}+U$是过(π,e)且斜率$k=2$的直线；

R^2/U是R^2中所有斜率$k=2$的直线的集合.

- 商空间的基与维数.

商映射$\pi: V\to V/U$，$\pi(\vec{v})=\vec{v}+U$.

易证π是线性映射，即$\pi\in\mathcal{L}(V,V/U)$.

显然π是满射，即$\mathrm{Ran}\,\pi=V/U$，且$\mathrm{Null}\,\pi=U$.

由线性映射基本定理，$\dim V=\dim U+\dim V/U$. 即：

- $\dim V/U=\dim V-\dim U$. (4-43)

- 设U是V的子空间，U的基为$\vec{u}_1,\vec{u}_2,\cdots,\vec{u}_m$，将其扩充为$V$的基.

$\vec{u}_1,\cdots,\vec{u}_m,\vec{v}_1,\cdots,\vec{v}_n$. 则$\vec{v}_1+U,\vec{v}_2+U,\cdots,\vec{v}_n+U$是$V/U$的基. (4-44)

证：设$\lambda_1,\lambda_2,\cdots,\lambda_n\in F$使得$0+U=\sum_{i=1}^{n}\lambda_i(\vec{v}_i+U)$

$\Rightarrow 0+U=\sum_{i=1}^{n}\lambda_i(\vec{v}_i+U)=\sum_{i=1}^{n}\lambda_i\vec{v}_i+U$.

$\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_n$线性无关 $\Rightarrow \lambda_i=0, \forall i \Rightarrow \vec{v}_1+U,\cdots,\vec{v}_n+U$线性无关.

且该线性无关组长度为 $n = \dim V - \dim U = \dim V/U$.

因此 $\vec{v}_1 + U, \cdots, \vec{v}_n + U$ 是 V/U 的基. Q.E.D.

- 映射 $\tilde{T}$.

对线性映射 $T \in \mathcal{L}(V, W)$,

定义 $\tilde{T}: V/\text{Null}\,T \to W$, $\tilde{T}(\vec{v} + \text{Null}\,T) = T(\vec{v})$

这里先证 $\tilde{T}$ 是良定义的.

即 $\forall \vec{v} \in V$, 假设 $\vec{v}' \in V$ 使得 $\vec{v} + \text{Null}\,T = \vec{v}' + \text{Null}\,T$.

须证 $T(\vec{v}) = T(\vec{v}')$:

$\vec{v} + \text{Null}\,T = \vec{v}' + \text{Null}\,T \Rightarrow \vec{v} - \vec{v}' \in \text{Null}\,T \Rightarrow T(\vec{v} - \vec{v}') = 0$

$\Rightarrow T(\vec{v}) = T(\vec{v}')$. Q.E.D.

$\tilde{T}$ 的性质:

i) $\tilde{T}$ 是线性映射

ii) $\tilde{T}$ 是单射

iii) $\text{Ran}\,\tilde{T} = \text{Ran}\,T$

iv) $V/\text{Null}\,T$ 与 $\text{Ran}\,T$ 同构 (※) (4-45)

证: i) $\forall \vec{v}, \vec{w} \in V$, $\lambda \in F$.

$\tilde{T}(\vec{v} + \text{Null}\,T) + \tilde{T}(\vec{w} + \text{Null}\,T) = T(\vec{v}) + T(\vec{w}) = T(\vec{v} + \vec{w})$

$= \tilde{T}[(\vec{v} + \vec{w}) + \text{Null}\,T] = \tilde{T}[(\vec{v} + \text{Null}\,T) + (\vec{w} + \text{Null}\,T)]$

$\tilde{T}[\lambda(\vec{v} + \text{Null}\,T)] = T(\lambda\vec{v}) = \lambda T(\vec{v}) = \lambda\tilde{T}(\vec{v} + \text{Null}\,T)$

$\Rightarrow \tilde{T} \in \mathcal{L}(V/\text{Null}\,T, W)$

ii) $\text{Null}\,\tilde{T} = \{\vec{v} + \text{Null}\,T \mid T(\vec{v}) = 0\} = \{\vec{v} + \text{Null}\,T \mid \vec{v} \in \text{Null}\,T\}$

$= \{0 + \text{Null}\,T\}$

$\Rightarrow \tilde{T}$ 是单射

iii) 显然.

iv) 若将 $\tilde{T}$ 看成 $V/\text{Null}\,T$ 到 $\text{Ran}\,T$ 上的映射, 则 iii) 表明 $\tilde{T}$ 是满射.

于是$\tilde{T}$是$V/\mathrm{Null}\,T$到$\mathrm{Ran}\,T$的同构，即$V/\mathrm{Null}\,T$同构于$\mathrm{Ran}\,T$.

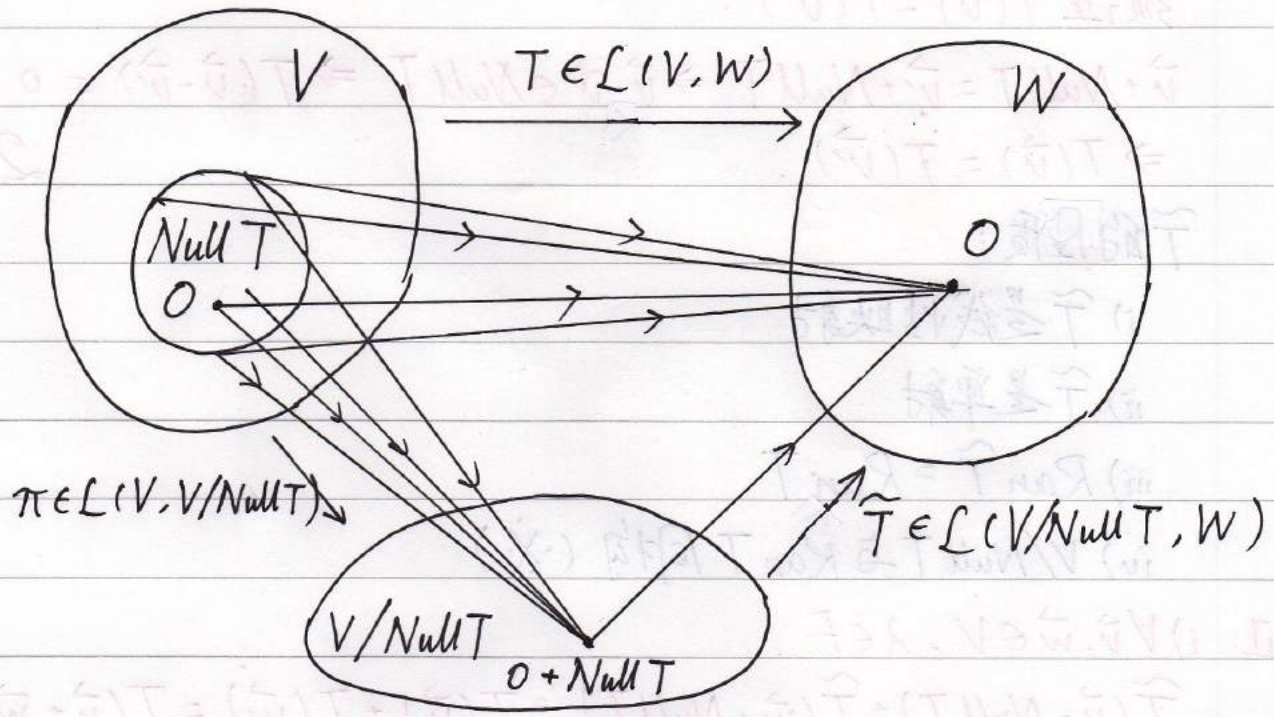
Q.E.D.

※：尽管由线性映射基本定理 $\dim \mathrm{Ran}\,T = \dim V - \dim \mathrm{Null}\,T$

以及商空间维数 $\dim V/\mathrm{Null}\,T = \dim V - \dim \mathrm{Null}\,T$

似乎也能得到$V/\mathrm{Null}\,T$同构于$\mathrm{Ran}\,T$，但这只限于V是有限维的，而定理(4-45)的证明不依赖维数。

即$V/\mathrm{Null}\,T$与$\mathrm{Ran}\,T$同构对无限维向量空间V仍成立。



谱论

5.1 本征值与本征矢

▶ 本章研究有限维向量空间中算子的结构. 找到线性算子(线性变换)的相似类中的最简矩阵——对角形或 Jordan 标准形.

▶ 不变子空间:

设 $T \in \mathcal{L}(V)$, 若 V 的子空间 U 有 $\forall \vec{u} \in U$, $T(\vec{u}) \in U$, 则称 U 是在 T 下的不变子空间.

- 一维不变子空间:

  $\forall \vec{v} \in V$, $span(\vec{v})$ 是 V 中的一维子空间.

  $span(\vec{v})$ 在 $T \in \mathcal{L}(V)$ 下不变, 当且仅当 $\exists \lambda \in F$, 使得 $T(\vec{v}) = \lambda \vec{v}$.

▶ 本征值、本征向量、本征空间

- 对 $T \in \mathcal{L}(V)$, 称数 $\lambda \in F$ 为 T 的本征值, 若 $\exists \vec{v} \in V - \{0\}$, 使得 $T(\vec{v}) = \lambda \vec{v}$. $\vec{v}$ 称为 T 相对应 λ 的本征向量.
- $$T(\vec{v}) = \lambda \vec{v} \iff (T - \lambda I)(\vec{v}) = 0.$$
- λ 是本征值 $\iff Null(T - \lambda I) \neq \{0\} \iff$ 算子 $T - \lambda I$ 非单射 $\iff$ 算子 $T - \lambda I$ 不可逆.
- 相对 λ 的本征向量张成的空间称为相对 λ 的本征空间, 记为 $E(\lambda, T) = \{\vec{v} \in V \mid T(\vec{v}) = \lambda \vec{v}\} = Null(T - \lambda I)$
- T 所有本征值构成的集合称为 T 的谱, 记为 $\sigma(T)$.
- 设 $T \in \mathcal{L}(V)$, $\lambda_1, \lambda_2, \cdots, \lambda_m$ 是 T 的互异本征值, $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_m$ 是对应的本征向量. 则 $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_m$ 线性无关. (5-1)

证: 假设 $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_m$ 线性相关.

设 k 是使 $\vec{v}_k = span(\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_{k-1})$ 成立的最小正整数, 即 $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_{k-1}$ 线性无关.

$\exists a_1, a_2, \cdots, a_{k-1} \in F$, 使得 $\vec{v}_k = \sum_{i=1}^{k-1} a_i \vec{v}_i$

$\Rightarrow T(\vec{v}_k) = T(\sum_{i=1}^{k-1} a_i\vec{v}_i) = \sum_{i=1}^{k-1} a_i T(\vec{v}_i)$

$\Rightarrow \lambda_k \vec{v}_k = \sum_{i=1}^{k-1} a_i \lambda_i \vec{v}_i$

由 $\lambda_k\vec{v}_k = \sum_{i=1}^{k-1} a_i\lambda_k\vec{v}_i$ 及上式得 $\sum_{i=1}^{k-1} a_i(\lambda_k-\lambda_i)\vec{v}_i = 0$

$\vec{v}_1,\cdots,\vec{v}_{k-1}$ 线性无关 $\Rightarrow a_1=\cdots=a_{k-1}=0 \Rightarrow \vec{v}_k = \sum_{i=1}^{k-1} a_i\vec{v}_i = 0$

这与 $\vec{v}_k$ 是本征向量矛盾．即 $\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_m$ 线性无关． Q.E.D

推论：有限维向量空间 V 上每个算子最多有 $\dim V$ 个互异本征值 (5-2)
(因为 V 中线性无关组长度不超过 $\dim V$)

▶ 求解本征值与本征向量

设 $T\in\mathcal{L}(V)$，T 在基 $\mathcal{B}=\{\vec{b}_1,\vec{b}_2,\cdots,\vec{b}_n\}$ 下的矩阵为 A，任意 $\vec{v}\in V$ 在基 $\mathcal{B}$ 下的坐标记为 $\vec{x}$.

※ 注意区分：$T\in\mathcal{L}(V)$． $A\in F^{n,n}$ (假设 $\dim V=n$)

$\vec{v}\in V$， $\vec{x}\in F^n$.

- 于是有 $\exists\,\vec{v}\in V-\{0\}$ 使得 $T(\vec{v})=\lambda\vec{v}$

$\Longleftrightarrow \exists\,\vec{x}\in F^n-\{0\}$ 使得 $A\vec{x}=\lambda\vec{x}$

$A=\{a_{i,j}\}$，$A\vec{x}=\lambda\vec{x}$ 写成 $(A-\lambda I)\vec{x}=0$

它有非零解 $\vec{x}$ 的充要条件为 $\det(\lambda I-A)=0$.

即 $$\begin{vmatrix} \lambda-a_{1,1} & -a_{1,2} & \cdots & -a_{1,n} \\ -a_{2,1} & \lambda-a_{2,2} & \cdots & -a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n,1} & -a_{n,2} & \cdots & \lambda-a_{n,n} \end{vmatrix} = 0.$$

$p(\lambda)=\det(\lambda I-A)$ 是关于 λ 的 n 次多项式，称为方阵 A 的特征多项式．$p(\lambda)$ 在 F 内的根就是 A 的本征值.

- 相似矩阵的特征多项式相同． (5-3)

证：设 $A,B\in F^{n,n}$．且 $\exists$ 可逆方阵 P，使得 $A=PBP^{-1}$

$\Rightarrow \lambda I-A = \lambda I-PBP^{-1} = \lambda PIP^{-1}-PBP^{-1} = P(\lambda I-B)P^{-1}$

$\det(\lambda I - A) = \det(P(\lambda I - B)P^{-1}) = \det P \cdot \det(\lambda I - B) \cdot \det(P^{-1})$

$\det P \cdot \det(P^{-1}) = 1 \Rightarrow \det(\lambda I - A) = \det(\lambda I - B)$. Q.E.D.

※ 该定理说明同一相似类有同样的特征多项式.

则算子的特征多项式可定义为在任意基下的矩阵的特征多项式.

• 求本征值 & 本征向量的算法归纳:

对方程 $T(\vec{v}) = \lambda\vec{v}$

i) 在 V 中取一组基, 求 T 在该基下的矩阵 A

ii) 计算特征多项式 $\det(\lambda I - A)$

iii) 求 $p(\lambda) = \det(\lambda I - A) = 0$ 的根 $\lambda_1, \lambda_2, \cdots, \lambda_m$

iv) 对每个 λ_i, 写出齐次方程组 $(A - \lambda_i I)\vec{x} = 0$.

v) 写出基础解系, 也即求出 $Null(A - \lambda_i I)$ 的基.

vi) 求出 v) 中基在坐标映射下的原像, 也即求出 $E(\lambda_i, T)$ 的基.

vii) 对应 λ_i 的特征向量就是 $E(\lambda_i, T)$ 中的任意非零向量.

举个栗子吧:

例 5-1: 设 V 中算子 T 在基 $\vec{b_1}, \vec{b_2}, \vec{b_3}$ 下的矩阵为

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$$

特征多项式 $p(\lambda) = \det(\lambda I - A)$

$$= \begin{vmatrix} \lambda-1 & -2 & -2 \\ -2 & \lambda-1 & -2 \\ -2 & -2 & \lambda-1 \end{vmatrix} = (\lambda-5)(\lambda+1)^2$$

$p(\lambda) = 0 \Rightarrow \lambda_1 = 5, \lambda_2 = -1$. 都是 T 的本征值.

对 $\lambda_1 = 5$, $A - \lambda_1 I$

$$= \begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \begin{cases} x_1 - x_3 = 0 \\ x_2 - x_3 = 0 \\ 0 = 0 \end{cases}$$

$$\Rightarrow \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathrm{span}(1,1,1)^T.$$

$\Rightarrow E(\lambda_1, T)$ 的基为 $\vec{b_1}+\vec{b_2}+\vec{b_3}$.

对 $\lambda_2 = -1$, $A - \lambda_2 I$

$$= \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \begin{cases} x_1 + x_2 + x_3 = 0 \\ 0 = 0 \\ 0 = 0 \end{cases}$$

$$\Rightarrow \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_2 - x_3 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

$$\in \mathrm{span}((-1,1,0)^T, (-1,0,1)^T)$$

$\Rightarrow E(\lambda_2, T)$ 的基为 $-\vec{b_1}+\vec{b_2}$, $-\vec{b_1}+\vec{b_3}$.

- 三角矩阵的本征值就是主对角线元素. (5-4)

证：对三角矩阵 $A \in F^{n,n}$, $A - \lambda I$ 也是三角矩阵，且

$\det(A-\lambda I) = (a_{1,1}-\lambda)(a_{2,2}-\lambda)\cdots(a_{n,n}-\lambda)$.

$p(\lambda) = 0 \Rightarrow \lambda_1 = a_{1,1}, \lambda_2 = a_{2,2}, \cdots, \lambda_n = a_{n,n}$. Q.E.D.

► 迹与行列式

- 相似矩阵的行列式相同. (5-5)

证：设 $A, B \in F^{n,n}$，且 ∃ 可逆方阵 P，使得 $A = PBP^{-1}$.

$\det A = \det(PBP^{-1}) = \det P \cdot \det B \cdot \det(P^{-1}) = \det B$. Q.E.D.

※ 同一相似类有相同的行列式.

算子的行列式可定义为在任意基下的矩阵的行列式.

- 矩阵的迹.

$\forall A \in F^{n,n}$, A的迹定义为主对角线元素之和.

$tr A = \sum_{i=1}^{n} a_{i,i}$.

- $\forall A, B \in F^{n,n}$, $tr(AB) = tr(BA)$ (5-6)

证: 令 $C = AB$. $c_{i,i} = \sum_{k=1}^{n} a_{i,k} b_{k,i}$

$\Rightarrow tr C = \sum c_{i,i} = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{i,k} b_{k,i}$

令 $D = BA \Rightarrow tr D = \sum_{i=1}^{n}\sum_{k=1}^{n} b_{i,k} a_{k,i} = \sum_{k=1}^{n}\sum_{i=1}^{n} a_{k,i} b_{i,k} = tr C$

$\Rightarrow tr(AB) = tr(BA)$ Q.E.D.

- 相似矩阵的迹相同. (5-7)

证: 设 $A, B \in F^{n,n}$, 且∃可逆方阵P, 使得 $A = PBP^{-1}$

$tr A = tr(PBP^{-1}) = tr(P \cdot BP^{-1}) = tr(BP^{-1} \cdot P) = tr B$. Q.E.D.

※ 同一相似类有相同的迹

算子的迹可定义为在任意基下矩阵的迹.

- 算子的迹与行列式.

设V上算子T的特征多项式$p(\lambda)$有C上的根$\lambda_1, \lambda_2, \cdots, \lambda_m$(计重根),

则有: i) $\det T = \prod_{i=1}^{m} \lambda_i$, ii) $tr T = \sum_{i=1}^{m} \lambda_i$ (5-8)

证: i) 由已知得 $\det(\lambda I - A) = \alpha(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_m)$.

$= \alpha\lambda^m - \alpha(\lambda_1 + \cdots + \lambda_m)\lambda^{m-1} + \cdots + (-1)^m \lambda_1 \cdots \lambda_m$

展开 $\det(\lambda I - A) = \sum_{\sigma \in perm(n)} (-1)^{r(\sigma)} a'_{\sigma_1,1} a'_{\sigma_2,2} \cdots a'_{\sigma_n,n}$

其中 $a'_{i,j} = \begin{cases} -a_{i,j}, & i \neq j \\ \lambda - a_{i,j}, & i = j \end{cases}$ (A是T在任意基下的矩阵)

求和式中λ的最高次项为 $\sigma_1 = 1, \sigma_2 = 2, \cdots, \sigma_n = n$ (标准次序)时取得

$r(\sigma) = 0$.

此项为 $(\lambda - a_{1,1})(\lambda - a_{2,2})\cdots(\lambda - a_{n,n}) = \lambda^n - (a_{1,1} + \cdots + a_{n,n})\lambda^{n-1} + \cdots$

此项与 $\det(\lambda I-A)$ 比较得 $\alpha=1$，$m=n$

即 $\det(\lambda I-A)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$

取 $p(0)=\det(-A)$ 代入上式：

$$\det A=(-1)^n\det(-A)=(-1)^n(-\lambda_1)(-\lambda_2)\cdots(-\lambda_n)$$
$$=\prod_{i=1}^{n}\lambda_i\ .$$

ii) 由 i) 中的展开式，有

$$\det(\lambda I-A)=(\lambda-a_{1,1})(\lambda-a_{2,2})\cdots(\lambda-a_{n,n})+q(\lambda)\ .$$

标值次序交换 1 次，则乘积减少两个含 λ 的因式，即 $q(\lambda)$ 最高只有 $(n-2)$ 次。

$$\Rightarrow\ \det(\lambda I-A)=\lambda^n-(a_{1,1}+a_{2,2}+\cdots+a_{n,n})\lambda^{n-1}+q(\lambda)$$

又有 $\det(\lambda I-A)=\lambda^n-(\lambda_1+\lambda_2+\cdots+\lambda_n)\lambda^{n-1}+\cdots+(-1)^n\lambda_1\lambda_2\cdots\lambda_m$

两式比较得 $a_{1,1}+a_{2,2}+\cdots+a_{n,n}=\lambda_1+\lambda_2+\cdots+\lambda_n$

即 $\operatorname{tr}A=\sum_{i=1}^{n}a_{i,i}=\sum_{i=1}^{n}\lambda_i$. Q.E.D.

※ 算子 T 的特征多项式可表示为

$$p(\lambda)=\lambda^n-\operatorname{tr}T\cdot\lambda^{n-1}+\cdots+(-1)^n\det T$$

▶ 本征值的重数

- 对多项式 $p(\lambda)$ 的根 λ_i，使得 $(\lambda-\lambda_i)^k$ 整除 $p(\lambda)$ 的最大整数 k 称为根 λ_i 的重数。
- 特征多项式的根 λ_i 的重数称为本征值 λ_i 的代数重数。
- 本征空间 $E(\lambda_i,T)$ 的维数称为本征值 λ_i 的几何重数。
- 复数域上的有限维向量空间 V 中的算子 T 的各本征值的代数重数之和等于 $\dim V$. (5-9)

证：该表述等价于 $p(\lambda)$ 在 $\mathbb{C}$ 上有 $\dim V$ 个根，而 $p(\lambda)$ 的次数为 $\dim V$。由代数基本定理，显然。 Q.E.D.

※ 该结论在实算子上不再成立，实算子甚至可能没有本征值。

例如矩阵 $A=\begin{pmatrix}1 & 2\\ -2 & 1\end{pmatrix}\in R^{2,2}$.

A 的特征多项式在 C 上有根 $\lambda_1=1+2i$，$\lambda_2=1-2i$.

而在 R 上无根. 即 $A\in R^{2,2}$ 无本征值.

- $\forall\lambda_i\in\sigma(T)$，其几何重数 $\leq$ 代数重数. (5-10)

证：设 T 是有限维向量空间 V 中的算子

设 $\vec{u}_1,\vec{u}_2,\cdots,\vec{u}_m$ 是 $E(\lambda_i,T)$ 的基，将其扩充为 V 的基 $\vec{u}_1,\cdots,\vec{u}_m,\vec{v}_1,\cdots,\vec{v}_n$. 考虑 T 在该基下的矩阵 A

$T(\vec{u}_1)=\lambda_i\vec{u}_1$，$T(\vec{u}_2)=\lambda_i\vec{u}_2$，$\cdots$，$T(\vec{u}_m)=\lambda_i\vec{u}_m$.

$$\text{于是 } A=\left[\begin{array}{c:c}\underbrace{\begin{matrix}\lambda_i & & 0\\ & \ddots & \\ 0 & & \lambda_i\end{matrix}}_{m\text{列}} & *\\ \hdashline 0 & *\end{array}\right]=\begin{pmatrix}\lambda_i I_m & B\\ 0 & C\end{pmatrix}$$

成分块上三角矩阵.

$$\Rightarrow p(\lambda)=\det(\lambda I-A)=\begin{vmatrix}(\lambda-\lambda_i)I_m & -B\\ 0 & \lambda I-C\end{vmatrix}$$

$$=\det[(\lambda-\lambda_i)I_m]\cdot\det(\lambda I-C)=(\lambda-\lambda_i)^m\det(\lambda I-C).$$

$p(\lambda)$ 被 $(\lambda-\lambda_i)^m$ 整除，则 λ_i 的代数重数 $\geq m=$ 几何重数. Q.E.D

- 代数重数将在广义本征空间中有更深的研究.

5.2 对角化

- 算子的对角化是找到一个基，使之成对角矩阵.

  矩阵的对角化是找到一个与之相似的对角矩阵.

- 矩阵的对角化可以方便地处理幂：

设 $A = PDP^{-1}$，$D = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$

则 $A^k = PD^kP^{-1}$，其中 $D^k = \mathrm{diag}(\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k)$

▶ 对角化算法.

若 $T \in \mathcal{L}(V)$ 在基 $\vec{v_1}, \vec{v_2}, \cdots, \vec{v_n}$ 下成对角矩阵 $\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$

则 $T(\vec{v_1}) = \lambda_1\vec{v_1}$，$T(\vec{v_2}) = \lambda_2\vec{v_2}$，$\cdots$，$T(\vec{v_n}) = \lambda_n\vec{v_n}$.

即 $\lambda_1, \lambda_2, \cdots, \lambda_n \in \sigma(T)$，$\vec{v_1}, \vec{v_2}, \cdots, \vec{v_n}$ 是对应 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 的本征向量（且它们线性无关）.

而另一方面，若 T 有线性无关的 $\dim V$ 个本征向量 $\vec{v_1}, \vec{v_2}, \cdots, \vec{v_n}$，则 T 在该基下的矩阵 A：

$T(\vec{v_1}) = \lambda_1\vec{v_1}$，$T(\vec{v_2}) = \lambda_2\vec{v_2}$，$\cdots$，$T(\vec{v_n}) = \lambda_n\vec{v_n}$

$\Rightarrow A = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$，且 $\lambda_1, \lambda_2, \cdots, \lambda_n \in \sigma(T)$.

• 可对角化：算子可对角化，若存在一个基使之成对角矩阵.

上面的讨论可总结为：

• 算子可对角化，当且仅当空间有一组由本征向量构成的基. (5-11)

▶ 可对角化的条件

• 本征值互异则可对角化.

若 $T \in \mathcal{L}(V)$ 有 $\dim V$ 个互异本征值，则 T 可对角化. (5-12)

证：设 $\lambda_1, \lambda_2, \cdots, \lambda_{\dim V} \in \sigma(T)$ 互异，则对应的本征向量 $\vec{v_1}, \vec{v_2}, \cdots, \vec{v}_{\dim V}$ 线性无关，并构成 V 的基. 由 (5-11)，T 可对角化. Q.E.D.

※ 然而并非所有算子都有足够的互异本征值. 那么在其它情况下，算子是否还能对角化呢？

接下来进入本征空间的讨论：

• 本征空间的和是直和. (5-13)

证：设 $T \in \mathcal{L}(V)$ 的本征空间为 $E(\lambda_1, T), E(\lambda_2, T), \cdots, E(\lambda_m, T)$.

$\forall \vec{v} \in E(\lambda_1, T) + E(\lambda_2, T) + \cdots + E(\lambda_m, T)$，$\vec{v} = \sum_{i=1}^{m} a_i\vec{v_i}$，$\vec{v_i} \in E(\lambda_i, T)$

属于不同本征空间的向量线性无关 $\Rightarrow a_1 = a_2 = \cdots = a_m = 0$

$\Rightarrow E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ 是直和. Q.E.D.

- 算子可对角化，当且仅当原空间可表示为本征空间的直和. (5-14)

证：充分性：若 $T \in \mathcal{L}(V)$，$V = \bigoplus_{i=1}^{m} E(\lambda_i, T)$.

在每个 $E(\lambda_i, T)$ 中取一组基，并成 V 中的一组基(定理(4-36))，该基由本征向量构成，于是 T 可对角化.

必要性：若 $T \in \mathcal{L}(V)$ 可对角化，则 V 中存在一组由本征向量构成的基，

即 $V \subset \bigoplus_{i=1}^{m} E(\lambda_i, T)$

又每个 $E(\lambda_i, T) \subset V \Rightarrow \bigoplus_{i=1}^{m} E(\lambda_i, T) \subset V$

$\Rightarrow V = \bigoplus_{i=1}^{m} E(\lambda_i, T)$ Q.E.D.

※ 算子可对角化，当且仅当 $\dim V = \sum_{i=1}^{m} \dim E(\lambda_i, T)$ (5-15)

证：由定理4-38，$\dim\left(\bigoplus_{i=1}^{m} E(\lambda_i, T)\right) = \sum_{i=1}^{m} \dim E(\lambda_i, T)$.

算子可对角化 $\Leftrightarrow V = \bigoplus_{i=1}^{m} E(\lambda_i, T) \Leftrightarrow \dim V = \dim\left(\bigoplus_{i=1}^{m} E(\lambda_i, T)\right)$

(由于 $\bigoplus_{i=1}^{m} E(\lambda_i, T) \subset V$) $\Leftrightarrow \dim V = \sum_{i=1}^{m} \dim E(\lambda_i, T)$ Q.E.D.

- $\mathbb{C}$ 上算子可对角化当且仅当 $\forall \lambda_i \in \sigma(T)$，其代数重数 = 几何重数. (5-16)

证：$\forall \lambda_i \in \sigma(T)$，记代数重数为 d_i，几何重数为 $\dim E(\lambda_i, T)$.

由定理(5-9)，$\dim V = \sum d_i$.

算子可对角化 $\Leftrightarrow \dim V = \sum \dim E(\lambda_i, T)$.

由 $\forall \lambda_i \in \sigma(T)$，$d_i \geq \dim E(\lambda_i, T)$ (定理(5-10))

~ 等价于 $\forall \lambda_i \in \sigma(T)$，$d_i = \dim E(\lambda_i, T)$. Q.E.D.

- 所有 $T \in \mathcal{L}(V)$ 可对角化的等价表述总结：

i) V 中存在由 T 的本征向量构成的基.

ii) $V = \bigoplus_{i} E(\lambda_i, T)$

iii) $\dim V = \sum_{i} \dim E(\lambda_i, T)$

iv) $\forall \lambda_i \in \sigma(T)$，代数重数 = 几何重数 (要求 V 定义在 $\mathbb{C}$ 上)

▸ 对角化算法

设 $T\in\mathcal{L}(V)$ 可对角化，T 的本征值为 $\lambda_1,\lambda_2,\cdots,\lambda_m$，重数为 $d_1,d_2,\cdots,d_m$

则可在每个 $E(\lambda_i,T)$ 中取一个基 $\vec{v}_{i,1},\vec{v}_{i,2},\cdots,\vec{v}_{i,d_i}$，并成 V 的基：

$\vec{v}_{1,1},\cdots,\vec{v}_{1,d_1},\cdots,\vec{v}_{m,1},\cdots,\vec{v}_{m,d_m}$.

在该基下，T 成对角矩阵

$$D=\operatorname{diag}(\underbrace{\lambda_1,\cdots,\lambda_1}_{d_1\text{个}\lambda_1},\cdots,\underbrace{\lambda_i,\cdots,\lambda_i}_{d_i\text{个}\lambda_i},\cdots,\underbrace{\lambda_m,\cdots,\lambda_m}_{d_m\text{个}\lambda_m}).$$

例 5-2：对角化矩阵 $A=\begin{pmatrix}1&2\\-2&1\end{pmatrix}$

特征多项式 $\det(\lambda I-A)$

$$=\begin{vmatrix}\lambda-1&-2\\2&\lambda-1\end{vmatrix}=(\lambda-1)^2+2^2$$

$\det(\lambda I-A)=0\Rightarrow\lambda_1=1+2i$，$\lambda_2=1-2i$.

对 $\lambda_1=1+2i$，$A-\lambda_1 I$

$$=\begin{pmatrix}-2i&2\\-2&-2i\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1&i\\0&0\end{pmatrix}\Rightarrow\begin{cases}x_1+ix_2=0\\0=0\end{cases}$$

$$\Rightarrow\vec{x}=\begin{pmatrix}-ix_2\\x_2\end{pmatrix}=x_2\begin{pmatrix}-i\\1\end{pmatrix}，取\ x_2=1$$

对 $\lambda_2=1-2i$，$A-\lambda_2 I$

$$=\begin{pmatrix}2i&2\\-2&2i\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1&-i\\0&0\end{pmatrix}\Rightarrow\begin{cases}x_1-ix_2=0\\0=0\end{cases}$$

$$\Rightarrow\vec{x}=\begin{pmatrix}ix_2\\x_2\end{pmatrix}=x_2\begin{pmatrix}i\\1\end{pmatrix}，取\ x_2=1.$$

$$\Rightarrow A=\begin{pmatrix}-i&i\\1&1\end{pmatrix}\begin{pmatrix}1+2i&0\\0&1-2i\end{pmatrix}\begin{pmatrix}-i&i\\1&1\end{pmatrix}^{-1}=\begin{pmatrix}-i&i\\1&1\end{pmatrix}\begin{pmatrix}1+2i&0\\0&1-2i\end{pmatrix}\begin{pmatrix}1&-i\\-1&-i\end{pmatrix}$$

例5-3 Fibonacci数列.

$a_1=1$, $a_2=1$, $a_n=a_{n-1}+a_{n-2}$, 求$\{a_n\}$.

解: $\begin{cases} a_{n+1}=a_n+a_{n-1} \\ a_n=a_n \end{cases} \Rightarrow \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix}$

$$\Rightarrow \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^2 \begin{pmatrix} a_{n-1} \\ a_{n-2} \end{pmatrix} = \cdots = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{n-1} \begin{pmatrix} a_2 \\ a_1 \end{pmatrix}$$

为了求$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{n-1}$, 将其对角化:

令$A=\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $\det(\lambda I-A)=\begin{vmatrix} \lambda-1 & -1 \\ -1 & \lambda \end{vmatrix}=\lambda^2-\lambda-1$

$\det(\lambda I-A)=0 \Rightarrow \lambda_1=\frac{1+\sqrt5}{2}$, $\lambda_2=\frac{1-\sqrt5}{2}$

对λ_1, $A-\lambda_1 I$

$$=\begin{pmatrix} 1-\lambda_1 & 1 \\ 1 & -\lambda_1 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & -\lambda_1 \\ 0 & 0 \end{pmatrix} \Rightarrow \begin{cases} x_1-\lambda_1 x_2=0 \\ 0=0 \end{cases} \quad (\text{已用}\lambda_1^2=\lambda_1+1)$$

取$\vec{x}=(\lambda_1, 1)^T$

对λ_2, $A-\lambda_2 I$

$$=\begin{pmatrix} 1-\lambda_2 & 1 \\ 1 & -\lambda_2 \end{pmatrix} \sim \begin{pmatrix} 1 & -\lambda_2 \\ 0 & 0 \end{pmatrix}.\ \text{同理取}\vec{x}=(\lambda_2, 1)^T$$

$$\Rightarrow A=\begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}^{-1}$$

$$\Rightarrow A^{n-1}=\frac{1}{\lambda_1-\lambda_2}\begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1^{n-1} & 0 \\ 0 & \lambda_2^{n-1} \end{pmatrix} \begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix} = \frac{1}{\sqrt5}\begin{pmatrix} \lambda_1^n-\lambda_2^n & -\lambda_2\lambda_1^n+\lambda_1\lambda_2^n \\ \lambda_1^{n-1}-\lambda_2^{n-1} & -\lambda_2\lambda_1^{n-1}+\lambda_1\lambda_2^{n-1} \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \frac{1}{\sqrt5}\begin{pmatrix} \lambda_1^n-\lambda_2^n & -\lambda_2\lambda_1^n+\lambda_1\lambda_2^n \\ \lambda_1^{n-1}-\lambda_2^{n-1} & -\lambda_2\lambda_1^{n-1}+\lambda_1\lambda_2^{n-1} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\Rightarrow a_n=\frac{1}{\sqrt5}[\lambda_1^{n-1}(1-\lambda_2)-\lambda_2^{n-1}(1-\lambda_1)]=\frac{1}{\sqrt5}[(\frac{1+\sqrt5}{2})^n-(\frac{1-\sqrt5}{2})^n]$$

5.3 不变子空间

- 重新抄写不变子空间的定义：

  设 $T \in \mathcal{L}(V)$，若 V 的子空间 U 有 $\forall \vec{u} \in U$，$T(\vec{u}) \in U$。则称 U 是 V 在 T 下的不变子空间。

- 研究算子 $T \in \mathcal{L}(V)$ 的性质，若能将 T 的定义域限制在 V 的某个子空间 U_i 上，维数更低因而更方便。然而 V 上的算子 T 未必是 U_i 上的算子，这使一些操作（如取幂）变得困难。

  若考虑 V 上的不变子空间 U_i，则 V 上算子 T 同时也是 U_i 上的算子。

  于是目标就是找到 V 的直和分解：$V = \bigoplus_{i=1}^{m} U_i$，且每个 U_i 都是不变子空间。

▶ 不变子空间上的算子。

设 $T \in \mathcal{L}(V)$，U 是 V 在 T 下的不变子空间。

限制算子 $T|_U \in \mathcal{L}(U)$，$T|_U(\vec{u}) = T(\vec{u})$

即把 T 的定义域限定在 U 上。

商算子（某些国内教材叫诱导变换）

$T/U \in \mathcal{L}(V/U)$，$(T/U)(\vec{v}+U) = T(\vec{v})+U$

这里先证 T/U 是良定义的。

即 $\forall \vec{v}+U \in V/U$，假设 $\vec{w} \in V$ 使得 $\vec{w}+U = \vec{v}+U$。

须证 $T(\vec{v})+U = T(\vec{w})+U$：

$\vec{v}+U = \vec{w}+U \Rightarrow \vec{v}-\vec{w} \in U$。$U$ 是不变子空间 $\Rightarrow T(\vec{v}-\vec{w}) \in U$。

$\Rightarrow T(\vec{v})+U = T(\vec{w})+U$。 Q.E.D.

▶ T 的典型子空间。

- 显然 $\{0\}$ 与 V 是 $T \in \mathcal{L}(V)$ 的不变子空间，称为平凡不变子空间。
- $\operatorname{Null} T$ 是 T 的不变子空间。 (5-17)

证：$\forall \vec{v} \in \operatorname{Null} T$，$T(\vec{v}) = 0 \in \operatorname{Null} T$。 Q.E.D.

- RanT是T的不变子空间. (5-18)

证: $\forall \vec{v} \in RanT$, $T(\vec{v}) \in RanT$ Q.E.D.

▸ 一维不变子空间

- T的本征向量张成一维不变子空间

  设 $\vec{v}$ 是对应 $\lambda \in \sigma(T)$ 的本征向量.

  $\forall \vec{w} \in span(\vec{v})$, $T(\vec{w}) = \lambda \vec{w} \in span(\vec{v})$

- $T \in \mathcal{L}(V)$ 可对角化, 当且仅当V可表示为一维不变子空间的直和.

  $V = \bigoplus_{i=1}^{n} span(\vec{v}_i)$ (5-19)

(该表述显然与定理(5-11)等价)

▸ 并非所有算子都可对角化, 对一些不具有对角形的算子, 下面定理给出了具有准对角形的条件.

- $T \in \mathcal{L}(V)$ 在某个基下成准对角形, 当且仅当V能分解成不变子空间的直和. (5-20)

证: 充分性: 若 $V = \bigoplus_{i=1}^{m} U_i$, 其中 U_i 是T的 n_i 维不变子空间.

在每个 U_i 中取一组基 $\vec{u}_{i,1}, \vec{u}_{i,2}, \cdots, \vec{u}_{i,n_i}$, 并成V的一组基.

考虑T在该基下的矩阵A.

由于 $T(\vec{u}_{i,1}), \cdots, T(\vec{u}_{i,n_i}) \in span(\vec{u}_{i,1}, \cdots, \vec{u}_{i,n_i})$.

$\Rightarrow A = diag(A_1, A_2, \cdots, A_m)$, 其中 $A_i \in F^{n_i, n_i}$.

必要性: 若 $A = diag(A_1, A_2, \cdots, A_m)$.

把基分成m段: $\vec{u}_{1,1}, \cdots, \vec{u}_{1,n_1}, \cdots, \vec{u}_{m,1}, \cdots, \vec{u}_{m,n_m}$, 其中 n_i 是 A_i 的阶.

令 $U_i = span(\vec{u}_{i,1}, \vec{u}_{i,2}, \cdots, \vec{u}_{i,n_i})$, 于是 $T(\vec{u}_{i,1}), \cdots, T(\vec{u}_{i,n_i}) \in span(\vec{u}_{i,1}, \cdots, \vec{u}_{i,n_i})$ $\Rightarrow$ U_i 是T的不变子空间. $V = \sum_{i=1}^{m} U_i$

且 $\dim \sum_{i=1}^{m} U_i = \dim V = \sum_{i=1}^{m} n_i = \sum_{i=1}^{m} \dim U_i$

$\Rightarrow V = \bigoplus_{i=1}^{m} U_i$. Q.E.D.

- 设 $T \in \mathcal{L}(V)$ 可对角化, 则 $\forall$ 不变子空间U, $T|_U$ 也可对角化. (5-21)

证: $T \in \mathcal{L}(V)$ 可对角化 $\Rightarrow V = \bigoplus_{i=1} E(\lambda_i, T)$

令 $W_i = U \cap E(\lambda_i, T)$, $i = 1, \cdots, m$.

现证 $U = \bigoplus_{i=1}^{m} W_i$:

i) $\sum_{i=1}^{m} W_i$ 是直和.

考虑 $0 = \sum_{i=1}^{m} \vec{v_i}$, 其中 $\vec{v_i} \in W_i \subseteq E(\lambda_i, T)$.

由于 $\sum_{i=1}^{m} E(\lambda_i, T)$ 是直和, $\vec{v_1} = \vec{v_2} = \cdots = \vec{v_m} = 0 \Rightarrow \sum_{i=1}^{m} W_i$ 是直和

ii) $U = \sum_{i=1}^{m} W_i$:

$W_1, W_2, \cdots, W_m \subset U \Rightarrow \sum_{i=1}^{m} W_i \subset U$.

$\forall \vec{v} \in U \subseteq V$, $\vec{v} = \sum_{i=1}^{m} \vec{v_i}$ $(\vec{v_i} \in E(\lambda_i, T))$

$T(\vec{v}) = \sum_{i=1}^{m} \lambda_i \vec{v_i}$

$\vdots$

$T^{m-1}(\vec{v}) = \sum_{i=1}^{m} \lambda_i^{m-1} \vec{v_i}$

$$\begin{pmatrix} \vec{v} \\ T(\vec{v}) \\ \vdots \\ T^{m-1}(\vec{v}) \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_m \\ \vdots & \vdots & & \vdots \\ \lambda_1^{m-1} & \lambda_2^{m-1} & \cdots & \lambda_m^{m-1} \end{pmatrix} \begin{pmatrix} \vec{v_1} \\ \vec{v_2} \\ \vdots \\ \vec{v_m} \end{pmatrix}$$

右端的 m 阶方阵的行列式为 Vandermonde 行列式, 因此方阵可逆.

$\Rightarrow \operatorname{span}(\vec{v}, T(\vec{v}), \cdots, T^{m-1}(\vec{v})) = \operatorname{span}(\vec{v_1}, \vec{v_2}, \cdots, \vec{v_m})$

U 是 T 的不变子空间. $\vec{v}, T(\vec{v}), \cdots, T^{m-1}(\vec{v}) \in U$.

$\Rightarrow \vec{v_1}, \vec{v_2}, \cdots, \vec{v_m} \in \operatorname{span}(\vec{v_1}, \vec{v_2}, \cdots, \vec{v_n}) = \operatorname{span}(\vec{v}, T(\vec{v}), \cdots, T^{m-1}(\vec{v})) \subset U$.

$\Rightarrow \vec{v_i} \in U \cap E(\lambda_i, T) = W_i$

$\vec{v} = \sum_{i=1}^{m} \vec{v_i} \in \sum_{i=1}^{m} W_i \Rightarrow U \subset \sum_{i=1}^{m} W_i$

$\Rightarrow U = \sum_{i=1}^{m} W_i$.

综上 $U = \bigoplus_{i=1}^{m} W_i$, 在每个 W_i 中取一组基并成 U 的基, 该基由 T 的本征向量构成. 因此 $T|_U$ 可对角化. Q.E.D.

▸ 上三角矩阵

- $\mathbb{C}$上的算子都有上三角矩阵

设V是$\mathbb{C}$上的有限维向量空间. $\forall T\in\mathcal{L}(V)$, V上存在一个基使得T成上三角形. (5-22)

证: 借助商映射, 对V的维数作归纳法.

i) $\dim V=1$ 时结论显然成立

ii) 假设 $\dim V=n$ 时, T有上三角形

iii) 当 $\dim V=n+1$ 时, 设$\vec{v}_0$是T的一个本征向量.

记 $U=\mathrm{span}(\vec{v}_0)\Rightarrow \dim V/U=n$.

由ii), V/U上有基$\vec{v}_1+U, \vec{v}_2+U, \cdots, \vec{v}_n+U$. 使得$T/U\in\mathcal{L}(V/U)$成上三角形. 且由上三角矩阵的性质, 有:

$(T/U)(\vec{v}_i+U)\in\mathrm{span}(\vec{v}_1+U,\cdots,\vec{v}_i+U)$. $i=1,2,\cdots,n$.

$\Rightarrow T(\vec{v}_i)\in\mathrm{span}(\vec{v}_0,\vec{v}_1,\cdots,\vec{v}_i)$, $i=1,2,\cdots,n$.

又$T(\vec{v}_0)\in\mathrm{span}(\vec{v}_0)\Rightarrow T$在$V$的基$\vec{v}_0,\vec{v}_1,\cdots,\vec{v}_n$下成上三角形. Q.E.D

※ 上三角矩阵性质归纳:

若$T\in\mathcal{L}(V)$在基$\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_n$下成上三角矩阵 $A=\{a_{i,j}\}_{i,j=1}^{n}$

i) $\forall i=1,2,\cdots,n$, $T\vec{v}_i\in\mathrm{span}(\vec{v}_1,\cdots,\vec{v}_i)$

ii) $\forall i=1,2,\cdots,n$, $\mathrm{span}(\vec{v}_1,\cdots,\vec{v}_i)$是$T$的不变子空间

iii) $\det T=\det A=a_{1,1}a_{2,2}\cdots a_{n,n}$

iv) $\sigma(T)=\{a_{1,1},a_{2,2},\cdots,a_{n,n}\}$

v) T可逆当且仅当$\forall a_{i,i}\neq 0$.

- 接下来几节的讨论会主要集中在复向量空间上, 因其特殊的性质:

i) 代数基本定理 (任意非常数的复系数多项式都有零点)

$\Rightarrow$ ii) 复向量空间上的算子都有本征值, 且代数重数和为$\dim V$.

$\Rightarrow$ iii) 复向量空间上的算子都有上三角矩阵

5.4 零化多项式

- 这节将涉及非常大量的多项式理论，这部分的系统证明将放入5.*更多注解中.

▶ Some motivations.

- 算子的最简表示是对角形，对应空间的本征空间分解.
- 有些算子无法对角化，那么可以退一步求出准对角形表示，即对应空间的不变子空间分解. 找到不变子空间 $W_1, W_2, \cdots, W_m$，使 $V = \bigoplus_{i=1}^{m} W_i$.
- 找到不变子空间直和分解，需要借助多项式的因式分解.

▶ 算子的多项式

- 设 $T \in \mathcal{L}(V)$，$m \in \mathbb{Z}^+$.

  定义算子的幂 $T^m = \underbrace{T \cdots T}_{m个}$

- 规定 T^0 为 V 上恒等算子 I.
- $T^{m+n} = T^m \circ T^n$，$T^{mn} = (T^m)^n = (T^n)^m$
- $p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m$ 是 V 上的算子，对应的多项式为 $p(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_m\lambda^m$.
- 算子的多项式一定是可交换的：$p(T)q(T) = q(T)p(T)$
- 算子 T 的全体多项式的集合记为 $F[T]$. $F[T]$ 是交换环.

▶ 算子多项式的零空间.

- $\forall p \in F[T]$，$\mathrm{Null}\, p(T)$ 是 T 的不变子空间. (5-23)

证：$\forall \vec{v} \in \mathrm{Null}\, p(T)$. $p(T)(\vec{v}) = 0$.

$\Rightarrow p(T)[T(\vec{v})] = T \cdot p(T)(\vec{v}) = T(0) = 0$

$\Rightarrow T(\vec{v}) \in \mathrm{Null}\, p(T)$. $\mathrm{Null}\, p(T)$ 在 T 下不变. Q.E.D.

- 设 $p(\lambda) \in F[\lambda]$ 在 F 内的因式分解为 $p(\lambda) = p_1(\lambda)p_2(\lambda)$.（即 p_1 与 p_2 互素），则 $\mathrm{Null}\, p(T) = \mathrm{Null}\, p_1(T) \oplus \mathrm{Null}\, p_2(T)$ (5-24)

证: i) 先证 $\text{Null}\,p(T) \supset \text{Null}\,p_1(T) + \text{Null}\,p_2(T)$:

$\forall \vec{v} \in \text{Null}\,p_1(T)$, $p_1(T)(\vec{v}) = 0$

$\Rightarrow p(T)(\vec{v}) = p_2(T)p_1(T)(\vec{v}) = 0 \Rightarrow \vec{v} \in \text{Null}\,p(T)$

$\Rightarrow \text{Null}\,p_1(T) \subset \text{Null}\,p(T)$ 同理 $\text{Null}\,p_2(T) \subset \text{Null}\,p(T)$

$\Rightarrow \text{Null}\,p_1(T) + \text{Null}\,p_2(T) \subset \text{Null}\,p(T)$

ii) 证 $\text{Null}\,p(T) \subset \text{Null}\,p_1(T) + \text{Null}\,p_2(T)$

$p_1 \cdot p_2$ 互素 $\Rightarrow \forall\, r(\lambda), s(\lambda) \in F[\lambda]$

s.t. $r(\lambda)p_1(\lambda) + s(\lambda)p_2(\lambda) = 1$ (定理5-

$\Rightarrow r(T)p_1(T) + s(T)p_2(T) = I$

$\forall \vec{v} \in V$, $\vec{v} = I(\vec{v}) = r(T)p_1(T)(\vec{v}) + s(T)p_2(T)(\vec{v})$

记 $\vec{v}_1 = s(T)p_2(T)(\vec{v})$, $\vec{v}_2 = r(T)p_1(T)(\vec{v})$.

若 $\vec{v}_1 \in \text{Null}\,p(T)$, $p_1(T)(\vec{v}_1) = p_1(T)s(T)p_2(T)(\vec{v})$

$= p(T)s(T)(\vec{v}) = 0 \Rightarrow \vec{v}_1 \in \text{Null}\,p_1(T)$

同理, 若 $\vec{v}_2 \in \text{Null}\,p(T)$, $\vec{v}_2 \in \text{Null}\,p_2(T)$.

$\vec{v} = \vec{v}_1 + \vec{v}_2 \Rightarrow \text{Null}\,p(T) \subset \text{Null}\,p_1(T) + \text{Null}\,p_2(T)$

iii) 证 $\text{Null}\,p_1(T)$ 与 $\text{Null}\,p_2(T)$ 是直和 .

$\forall \vec{v} \in \text{Null}\,p_1(T) \cap \text{Null}\,p_2(T)$.

$\vec{v} = r(T)p_1(T)(\vec{v}) + s(T)p_2(T)(\vec{v}) = 0 + 0 = 0$

$\Rightarrow \text{Null}\,p_1(T) \cap \text{Null}\,p_2(T) = \{0\} \Rightarrow \text{Null}\,p_1(T) + \text{Null}\,p_2(T)$ 是直和

由 i), ii), iii), $\text{Null}\,p(T) = \text{Null}\,p_1(T) \oplus \text{Null}\,p_2(T)$ Q.E.D

※ 推论: 把以上定理用数学归纳法推广,

若 $p(\lambda) \in F[\lambda]$ 有因式分解 $p(\lambda) = p_1(\lambda)p_2(\lambda)\cdots p_s(\lambda)$

则 F 上算子 T 满足 $\text{Null}\,p(T) = \bigoplus_{i=1}^{s} \text{Null}\,p_i(T)$ (5-25)

- 每给一个多项式 $p(\lambda)$, 定理5-25就给出了 $\text{Null}\,p(T)$ 的一个不变子空间直和分解 .

- 最终目标是找到 V 的不变子空间直和分解.

$\Rightarrow$ 找到一类 $p(\lambda)$, s.t. $Null\, p(T) = V$

$\Rightarrow \forall \vec{v} \in V, p(T)(\vec{v}) = 0 \Rightarrow p(T) = 0$.

也即目标是找到 $p(\lambda)$ 使 $p(T)$ 成为零算子

- 这样的 $p(\lambda)$ 称为 T 的零化多项式.
- 目标变为找 T 的零化多项式. 以下定理给出了一种最简单的方法:

▶ Cayley-Hamilton 定理

- C 上算子 T 的特征多项式是零化多项式 (5-26)

证: 设特征多项式 $p(\lambda) = \det(\lambda I - T)$.

伪证: $p(T) = \det(TI - T) = \det 0 = 0$ (why?)

正常证明:

设 $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n$ 是 V 的基, T 在该基下成上三角矩阵 A.

设 A 的主对角线元素为 $\lambda_1, \lambda_2, \cdots, \lambda_n$

(也即计重数的全体本征值).

$\Rightarrow p(\lambda) = (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$

$\Rightarrow p(T) = (T-\lambda_1 I)(T-\lambda_2 I)\cdots(T-\lambda_n I)$

A 是上三角矩阵 $\Rightarrow T(\vec{v}_i) = \lambda_i \vec{v}_i + \sum_{j=1}^{i-1} a_j \vec{v}_j \in span(\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_i)$

$\Rightarrow (T-\lambda_i I)(\vec{v}_i) = T(\vec{v}_i) - \lambda_i \vec{v}_i = \sum_{j=1}^{i-1} a_j \vec{v}_j \in span(\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_{i-1})$

为证明 $p(T) = 0$, 对 V 的维数作归纳法.

i) $n=1$ 时, $p(T) = T - \lambda_1 I$.

$(T-\lambda_1 I)(\vec{v}) \in span(0) = \{0\} \Rightarrow p(T) = 0$

ii) 假设 $n=k$ 时 $p(T) = 0$

iii) 当 $n=k+1$ 时, $p(T) = (T-\lambda_1 I)\cdots(T-\lambda_{k+1} I)$

由 ii), $\forall \vec{v} \in span(\vec{v}_1, \cdots, \vec{v}_k)$, $(T-\lambda_1 I)\cdots(T-\lambda_k I)(\vec{v}) = 0$.

$\forall \vec{v} \in V$, $(T-\lambda_{k+1} I)(\vec{v}) \in span(\vec{v}_1, \cdots, \vec{v}_k)$.

$\Rightarrow \forall \vec{v} \in V,\ (T-\lambda_1 I)\cdots(T-\lambda_k I)(T-\lambda_{k+1} I)(\vec{v}) = 0 \Rightarrow p(T) = 0.$

$\Rightarrow p(\lambda)$ 是T的零化多项式. Q.E.D.

※ 设T是复向量空间V上算子，互异本征值为 $\lambda_1, \lambda_2, \cdots, \lambda_m$，对应的代数重数为 $d_1, d_2, \cdots, d_m$.

特征多项式写成 $p(\lambda) = (\lambda-\lambda_1)^{d_1}(\lambda-\lambda_2)^{d_2}\cdots(\lambda-\lambda_m)^{d_m}$

$p(\lambda)$ 是零化多项式，由定理5-25，有：

$$V = Null\, p(T) = \bigoplus_{i=1}^{m} Null(T-\lambda_i I)^{d_i} \quad (5\text{-}27)$$

其中 $Null(T-\lambda_i I)^{d_i}$ 称为 λ_i 对应的广义本征空间（或"根子空间"），记为 $G(\lambda_i, T)$

$$\Rightarrow V = \bigoplus_{i=1}^{m} G(\lambda_i, T) \quad (5\text{-}28)$$

※ 这个直和分解将在下节中继续讨论：$T|_{G(\lambda_i,T)}$ 有怎么样的矩阵？

▶ 极小多项式

- 算子的极小多项式定义为次数最低的首一零化多项式（首一多项式是最高次项系数为1的多项式）
- 算子的极小多项式是唯一的. (5-29)

证：设 $m_1(\lambda)$ 与 $m_2(\lambda)$ 是T的极小多项式，假设 $m_1(\lambda) \neq m_2(\lambda)$

$m_1(T) = 0,\ m_2(T) = 0$

$\Rightarrow q(T) = m_1(T) - m_2(T) = 0$. 即 $q(\lambda) \in F[\lambda]$ 也是T的零化多项式且有 $deg(q) < deg(m_1) = deg(m_2)$.

q是比极小多项式次数更低的零化多项式，与定义矛盾.

$\Rightarrow m_1(\lambda) = m_2(\lambda)$. 极小多项式唯一. Q.E.D.

- 算子的任意零化多项式均可被极小多项式整除. (5-30)

证：设T的极小多项式为 $m(\lambda)$. $\forall q(\lambda) \in F[\lambda]$ s.t. $q(T) = 0$.

$\exists s(\lambda), r(\lambda) \in F[\lambda]$ s.t. $q(\lambda) = m(\lambda)s(\lambda) + r(\lambda)$, $deg(r) < deg(m)$.

（多项式的带余除法）.

$\Rightarrow q(T) = m(T)s(T) + r(T)$

$q(T) = 0$, $m(T) = 0 \Rightarrow r(T) = 0$.

若 $r(\lambda) \neq 0$，则 $r(\lambda)$ 是 T 的零化多项式（次数低于极小多项式），矛盾.

即 $r(\lambda) = 0 \Rightarrow m(\lambda) | q(\lambda)$. Q.E.D.

※ 推论：算子的特征多项式是极小多项式的多项式倍. (5-31)

· 极小多项式的零点恰为算子的全部本征值. (5-32)

证：设 T 的极小多项式为 $m(\lambda)$.

i) 若 λ_j 是 $m(\lambda)$ 的一个零点.

$\Rightarrow m(\lambda) = (\lambda - \lambda_j) q(\lambda)$, $\deg(q) < \deg(m)$.

$\Rightarrow m(T) = (T - \lambda_j I) q(T) = 0$

$\Rightarrow \forall \vec{v} \in V$, $(T - \lambda_j I)[q(T)(\vec{v})] = 0$.

$q(\lambda)$ 不是零化多项式. 即 $\exists \vec{v} \in V$ s.t. $q(T)(\vec{v}) = \vec{u} \neq 0$

$\Rightarrow (T - \lambda_j I)(\vec{u}) = 0 \Rightarrow \lambda_j \in \sigma(T)$.

ii) 若 $\lambda_j \in \sigma(T) \Rightarrow \exists \vec{v} \neq 0$ s.t. $T(\vec{v}) = \lambda_j \vec{v}$.

$\Rightarrow T[T(\vec{v})] = \lambda_j(\lambda_j \vec{v}) \Rightarrow T^2(\vec{v}) = \lambda_j^2(\vec{v})$.

推广得对正整数 k 都有 $T^k(\vec{v}) = \lambda_j^k(\vec{v})$.

$\forall \vec{v} \in V$, $m(T)(\vec{v}) = 0$. 设 $\deg(m) = m$

$\Rightarrow (T^m + \sum_{i=0}^{m-1} a_i T^i)(\vec{v}) = 0$.

$\Rightarrow (\lambda_j^m + \sum_{i=0}^{m-1} a_i \lambda_j^i)(\vec{v}) = 0 \Rightarrow m(\lambda_j)(\vec{v}) = 0 \Rightarrow m(\lambda_j) = 0$

$\Rightarrow \lambda_j$ 是 $m(\lambda)$ 的零点. Q.E.D.

· $\mathbb{C}$ 上算子可对角化，当且仅当其极小多项式 $m(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_m)$.
其中 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 为全部互异本征值. (5-33)

证：i) 若极小多项式 $m(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_m)$.

由定理 5-25，$V = \mathrm{Null}\, m(T) = \bigoplus_{i=1}^{m} \mathrm{Null}(T - \lambda_i I) = \bigoplus_{i=1}^{m} E(\lambda_i, T)$

算子可对角化.

ii) 若算子可对角化，则 V 中存在一组本征向量构成的基 $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n$.

$\forall i=1,2,\cdots,n.\ (T-\lambda_1 I)(T-\lambda_2 I)\cdots(T-\lambda_n I)(\vec{v}_i)=0$

$\forall \vec{v}\in V.\ \vec{v}=\sum_{i=1}^{n} a_i\vec{v}_i$

$(T-\lambda_1 I)(T-\lambda_2 I)\cdots(T-\lambda_n I)(\sum_{i=1}^{n} a_i\vec{v}_i)=0$

$\Rightarrow m(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$ 是零化多项式

假设存在零化多项式 $q(\lambda)$ s.t. $\deg(q)<\deg(m)=n$.

则 $q(\lambda)$ 的互异零点个数小于 n，$\exists \lambda_j\in\sigma(T)$ s.t. $q(\lambda_j)\neq 0$.

这与定理5-30 & 5-32矛盾. 这样的 $q(\lambda)$ 不存在.

即 $m(\lambda)$ 是极小多项式. Q.E.D.

5.5 幂零算子

- 设 N 是 $\mathbb{F}$ 上有限维向量空间 V 中的线性算子.

  若存在 $m\in\mathbb{N}^*$, s.t. $N^m=0$. 则称 N 是 V 中的幂零算子.

  最小的 m 称为 N 的幂零指数. 即 $N^{m-1}\neq 0$ 且 $N^m=0$.

- 幂零算子的唯一本征值等于 0. (5-34)

证：设 $\exists m\in\mathbb{N}^*$ s.t. $N^m=0 \Rightarrow q(\lambda)=\lambda^m$ 是零化多项式.

N 的极小多项式 $m(\lambda)$ 的零点即 $q(\lambda)$ 的唯一零点 0.

$\Rightarrow \sigma(N)=\{0\}$. Q.E.D.

▸ 循环不变子空间

- 设 N 是 V 中的幂零算子，幂零指数为 m.

  $\forall \vec{v}\in V$，若 $\exists k\in\mathbb{N}^*$ s.t. $N^{k-1}(\vec{v})\neq 0$ 且 $N^k(\vec{v})=0$. (显然 $k\le m$)

有 $N(\vec{v})=N(\vec{v})$ $\quad N[N(\vec{v})]=N^2(\vec{v})\ \cdots\ N[N^{k-2}(\vec{v})]=N^{k-1}(\vec{v})$

$N[N^{k-1}(\vec{v})]=N^k(\vec{v})=0$.

$\Rightarrow \mathrm{span}(\vec{v}, N(\vec{v}), \cdots, N^k(\vec{v}))$ 是 N 的不变子空间.

称为 $\vec{v}$ 张成的 N 的循环不变子空间，记为 $I(\vec{v})$.

- $\vec{v}, N(\vec{v}), \cdots, N^{k-1}(\vec{v})$ 是 $I(\vec{v})$ 的基，称为循环基 (5-35)

证：只须证线性无关.

假设 $N^{k-1}(\vec{v}), N^{k-2}(\vec{v}), \cdots, N^{0}(\vec{v})$ 线性相关.

$\exists j \in \{1, 2, \cdots, k\}$ s.t. $N^{k-j}(\vec{v}) \in \mathrm{span}(N^{k-1}(\vec{v}), N^{k-2}(\vec{v}), \cdots, N^{k-j+1}(\vec{v}))$

$\Rightarrow \exists \lambda_1, \lambda_2, \cdots, \lambda_{j-1} \in \mathbb{F}$. s.t. $N^{k-j}(\vec{v}) = \sum_{i=1}^{j-1} \lambda_i N^{k-j+i}(\vec{v})$

$\Rightarrow N^{j-1}[N^{k-j}(\vec{v})] = \sum_{i=1}^{j-1} \lambda_i N^{j-1}[N^{k-j+i}(\vec{v})] = \sum_{i=1}^{j-1} \lambda_i N^{k-1+i}(\vec{v})$

$\Rightarrow N^{k-1}(\vec{v}) = 0 \ \forall \vec{v} \in V. \Rightarrow N^{k-1} = 0$. 这与 $N^{k-1} \neq 0$ 矛盾. Q.E.D.

※ $N|_{I(\vec{v})}$ 在基 $\vec{v}, N(\vec{v}), \cdots, N^{m-1}(\vec{v})$ 下的矩阵

$$J_i = \begin{pmatrix} 0 & 1 & & & 0 \\ & 0 & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & 0 \end{pmatrix}$$

※ 对角线上方元素为 1，其余为 0.

J_i 称为一个 Jordan 块. (5-36)

左乘这个矩阵的效果：

$$\begin{pmatrix} 0 & 1 & & & 0 \\ & 0 & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & 0 \end{pmatrix} \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & \cdots & a_{3,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} = \begin{pmatrix} a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & \cdots & a_{3,n} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

即每行上移一行，最后一行变零，原第一行消失.

- 若向量空间可分解为循环不变子空间的直和，

即 $V = \bigoplus_{i=1}^{s} I(\vec{v}_i)$.

则幂零算子在 V 的某个基（循环不变子空间的基的并）下成准对角线. 若每个分块都是 Jordan 块，该准对角矩阵称为幂零算子

的Jordan形矩阵：　(5-37)

$$J=\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{pmatrix},\quad J_i=\begin{pmatrix} 0 & 1 & & 0 \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$$

下面完成直和分解的证明，此前要先证一个引理.

- 设N为V中非零幂零算子，N的本征空间$E(\lambda,N)$记为E（N只有本征值0）. 设$\vec{v}+E\neq E$，$I(\vec{v}+E)$为N/E在V/E内的k维循环不变子空间，则$I(\vec{v})$为N在V内的$(k+1)$维循环不变子空间，且$N^k(\vec{v})\in E$

证：$\dim I(\vec{v}+E)=k \Rightarrow (N/E)^k(\vec{v}+E)=E$　(5-38)

利用商映射$\pi\in\mathcal{L}(V,V/E)$，$\vec{v}\mapsto\vec{v}+E$.

$\pi[N^k(\vec{v})]=N^k(\vec{v})+E=(N^k/E)(\vec{v}+E)=(N/E)^k(\vec{v}+E)=E$

$\Rightarrow N^k(\vec{v})\in \mathrm{Null}\,\pi=E$

$\Rightarrow N^{k+1}(\vec{v})=N[N^k(\vec{v})]=0$　$(\forall\vec{v}\in E,\ N(\vec{v})=0)$.

假设$N^k(\vec{v})=0$.

$\Rightarrow N^k(\vec{v})=N[N^{k-1}(\vec{v})]=0 \Rightarrow N^{k-1}(\vec{v})\in E=\mathrm{Null}\,\pi$.

$\Rightarrow E=N^{k-1}(\vec{v})+E=(N/E)^{k-1}(\vec{v}+E)$

注意到$I(\vec{v}+E)=\mathrm{span}(\vec{v}+E,(N/E)(\vec{v}+E),\cdots,(N/E)^{k-1}(\vec{v}+E))$.

$\Rightarrow \dim I(\vec{v}+E)<k$ 与题设矛盾. $\Rightarrow N^k(\vec{v})\neq 0$.

即$I(\vec{v})=\mathrm{span}(\vec{v},N(\vec{v}),\cdots,N^k(\vec{v}))$是$k+1$维循环不变子空间.

Q.E.D.

- 设N为V中的非零幂零算子，则V可分解为N的循环不变子空间的直和$V=\bigoplus_{i=1}^{t}I(\vec{v}_i)$，其中$t=\dim E(0,N)$为本征值0的几何重数.　(5-39)

证：对V的维数作归纳法

$\dim V=1$时显然.

假设 $\dim V < n$ 时命题成立.

当 $\dim V = n$ 时, 记 N 的本征空间 $E(0, N) = E$.

商空间 V/E 的维数 $\dim V/E = \dim V - \dim E < \dim V$

由归纳假设, $V/E = \bigoplus_{i=1}^{s} I(\vec{v}_i + E)$.

其中 $\dim I(\vec{v}_i + E) = k_i$

由定理5-38, $I(\vec{v}_i)$ 为 N 在 V 内的 k_i+1 维循环子空间.

且 $I(\vec{v}_i) = \mathrm{span}(\vec{v}_i, N(\vec{v}_i), \cdots, N^{k_i}(\vec{v}_i))$, $N^{k_i}(\vec{v}_i) \in E$.

目标是证明 $V = I(\vec{v}_1) \oplus \cdots \oplus I(\vec{v}_s) \oplus I(\vec{u}_1) \oplus \cdots \oplus I(\vec{u}_r)$

i) 先证 $N^{k_1}(\vec{v}_1), N^{k_2}(\vec{v}_2), \cdots, N^{k_s}(\vec{v}_s)$ 线性无关:

$\exists \lambda_1, \lambda_2, \cdots, \lambda_s \in \mathbb{F}$, s.t. $\sum_{i=1}^{s} \lambda_i N^{k_i}(\vec{v}_i) = 0$.

显然 $k_i \geq 1$, 有 $N\left[\sum_{i=1}^{s} \lambda_i N^{k_i - 1}(\vec{v}_i)\right] = 0$.

$\Rightarrow \sum_{i=1}^{s} \lambda_i N^{k_i - 1}(\vec{v}_i) \in E$

利用商映射, $\pi\left[\sum_{i=1}^{s} \lambda_i N^{k_i-1}(\vec{v}_i)\right] = \sum_{i=1}^{s} \lambda_i \pi[N^{k_i-1}(\vec{v}_i)]$

$= \sum_{i=1}^{s} \lambda_i [N^{k_i-1}(\vec{v}_i) + E] = \sum_{i=1}^{s} \lambda_i (N/E)^{k_i-1}(\vec{v}_i + E) = E$

$(N/E)^{k_i-1}(\vec{v}_i + E) \in I(\vec{v}_i + E)$, 而 $I(\vec{v}_1 + E) + \cdots + I(\vec{v}_s + E)$ 是直和

$\Rightarrow (N/E)^{k_1-1}(\vec{v}_1 + E), \cdots, (N/E)^{k_s-1}(\vec{v}_s + E)$ 线性无关.

$\Rightarrow \lambda_1, \lambda_2, \cdots, \lambda_s = 0 \Rightarrow N^{k_1}(\vec{v}_1), \cdots, N^{k_s}(\vec{v}_s)$ 线性无关.

ii) 以上论证说明 $N^{k_1}(\vec{v}_1), \cdots, N^{k_s}(\vec{v}_s)$ 是 E 中的线性无关组,

扩充为 E 的基: $N^{k_1}(\vec{v}_1), \cdots, N^{k_s}(\vec{v}_s), \vec{u}_1, \cdots, \vec{u}_r$

令 $U = \mathrm{span}(\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_r)$. 要证 $V = I(\vec{v}_1) \oplus \cdots \oplus I(\vec{v}_s) \oplus U$:

iii) 证 $I(\vec{v}_1) + \cdots + I(\vec{v}_s) + U$ 是直和:

设 $0 = \vec{v}_1' + \cdots + \vec{v}_s' + \vec{u}$ $(\vec{v}_i' \in I(\vec{v}_i), \vec{u} \in U)$

$\Rightarrow E = \pi(0) = \pi(\vec{v}_1') + \cdots + \pi(\vec{v}_s') + \pi(\vec{u})$

$\vec{u} \in U \subset E \Rightarrow \pi(\vec{u}) = E \Rightarrow E = \sum_{i=1}^{s} \pi(\vec{v}_i'), \pi(\vec{v}_i') \in I(\vec{v}_i + E)$

$I(\vec{v}_1 + E) + \cdots + I(\vec{v}_s + E)$ 是直和 $\Rightarrow \pi(\vec{v}_i') = E$.

iv) 设 $\vec{v}_i' \in I(\vec{v}_i)$，$\vec{v}_i' = \sum_{j=0}^{k_i} \xi_j N^j(\vec{v}_i)$ $\quad(\xi_j \in F)$

$\pi(\vec{v}_i') = \sum_{j=0}^{k_i} \xi_j \pi[N^j(\vec{v}_i)] = \sum_{j=0}^{k_i-1} \xi_j (N/E)^j(\vec{v}_i + E) \in I(\vec{v}_i + E)$

（再次注意 $\pi[N^{k_i}(\vec{v}_i)] = N^{k_i}(\vec{v}_i) + E = E$）

则 $\pi(\vec{v}_i') = E \Rightarrow \sum_{j=0}^{k_i-1} \xi_j (N/E)^j(\vec{v}_i + E) = E$

由 $\vec{v}_i + E, (N/E)(\vec{v}_i + E), \cdots, (N/E)^{k_i-1}(\vec{v}_i + E)$ 是 $I(\vec{v}_i + E)$ 的基

得 $\xi_1 = \cdots = \xi_{k_i-1} = 0$

$\Rightarrow \vec{v}_i' = \xi_{k_i} N^{k_i}(\vec{v}_i)$

代入 $0 = \sum_{j=1}^{s} \vec{v}_j' + \vec{u}$ 得：$0 = \sum_{i=1}^{s} \xi_{k_i} N^{k_i}(\vec{v}_i) + \sum_{j=1}^{r} \zeta_j \vec{u}_j$

又由 $N^{k_1}(\vec{v}_1), \cdots, N^{k_s}(\vec{v}_s), \vec{u}_1, \cdots, \vec{u}_r$ 是 E 的基

$\Rightarrow \xi_{k_1} = \cdots = \xi_{k_s} = \zeta_1 = \cdots = \zeta_r = 0$

$\Rightarrow I(\vec{v}_1) + \cdots + I(\vec{v}_s) + U$ 是直和.

v) 通过计算直和的维数完成证明

$\dim(I(\vec{v}_1) \oplus \cdots \oplus I(\vec{v}_s) \oplus U)$

$= \sum_{i=1}^{s} \dim I(\vec{v}_i) + \dim U = \sum_{i=1}^{s} (k_i + 1) + r$

$= \sum_{i=1}^{s} k_i + (s + r) = \sum_{i=1}^{s} \dim I(\vec{v}_i + E) + \dim E$

$= \dim V/E + \dim E = \dim V$

$\Rightarrow V = I(\vec{v}_1) \oplus \cdots \oplus I(\vec{v}_s) \oplus U$.

vi) $U = \mathrm{span}(\vec{u}_1, \cdots, \vec{u}_r) \subset E$.

$\Rightarrow N(\vec{u}_1) = \cdots = N(\vec{u}_r) = 0$.

$\Rightarrow I(\vec{u}_j) = \mathrm{span}(\vec{u}_j)$ 是 N 的循环不变子空间.

$\Rightarrow U = \bigoplus_{j=1}^{r} \mathrm{span}(\vec{u}_j) = \bigoplus_{j=1}^{r} I(\vec{u}_j)$

$\Rightarrow V = I(\vec{v}_1) \oplus \cdots \oplus I(\vec{v}_s) \oplus I(\vec{u}_1) \oplus \cdots \oplus I(\vec{u}_r)$.

这里循环子空间的个数 $s + r = \dim E$，恰为本征值 0 的几何重数.

Q.E.D.

5.6 Jordan标准形

▸ 幂零算子的Jordan形

· 设N为V中的幂零算子，幂零指数为l. 存在V中的一组基，使得N在该基下的矩阵为Jordan形，且t阶Jordan块（$1\le t\le l$）的个数

$$n(t)=\operatorname{rank}N^{t-1}+\operatorname{rank}N^{t+1}-2\operatorname{rank}N^{t} \tag{5-40}$$

证：由5-39，V可分解为循环不变子空间的直和. 在每个循环不变子空间中取循环基，N在该空间的限制在循环基下的矩阵为Jordan块；循环不变子空间的循环基并成V的基，N在该基下成Jordan形矩阵.

下面考虑Jordan块的个数：

t阶Jordan块的秩 $\operatorname{rank}J_t=t-1$

且有 $\operatorname{rank}J_t^2=t-2$, …, $\operatorname{rank}J_t^{t-1}=1$, $\operatorname{rank}J_t^t=0$.

对整个Jordan形矩阵，有：

$\operatorname{rank}I=\sum_{i=1}^{l} i\cdot n(i)$, $\operatorname{rank}N=\sum_{i=1}^{l}\operatorname{rank}J_i\cdot n(i)=\sum_{i=2}^{l}(i-1)n(i)$

… $\operatorname{rank}N^t=\sum_{i=t+1}^{l}(i-t)n(i)$ $(t<l)$

$\operatorname{rank}N^{t-1}+\operatorname{rank}N^{t+1}-2\operatorname{rank}N^t$

$=\sum_{i=t}^{l}(i-t)n(i)+\sum_{i=t+2}^{l}(i-t)n(i)-2\sum_{i=t+1}^{l}(i-t)n(i)=n(t)$

易证$t=l-1$, l时上式仍满足. Q.E.D.

▸ 复向量空间上一般算子的Jordan形

· Jordan块 & Jordan形 (5-41)

Jordan块矩阵是形如以下的矩阵：

$$J_i=\begin{pmatrix}\lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1\\ & & & \lambda_i\end{pmatrix},\quad J=\begin{pmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s\end{pmatrix}$$

Jordan形矩阵是由Jordan块组成的对角矩阵.

- 若 $T-\lambda_i I$ 是 V 上的幂零算子，则存在 V 的一组基使得 T 的矩阵成 Jordan 形. 该基称为 Jordan 基. (5-42)

证：由 5-40，V 上存在一组基使得 $T-\lambda_i I$ 的矩阵成 Jordan 形，且在该基下 $\lambda_i I$ 的矩阵为 $diag(\lambda_i, \cdots, \lambda_i)$，于是 $T=(T-\lambda_i I)+\lambda_i I$ 在该基下的矩阵是形如 5-41 的 Jordan 形. Q.E.D.

- 设 T 是有限维复向量空间 V 上的算子，存在 V 的一组基，使得 T 在该基下成 Jordan 形矩阵. (5-43)

证：设 T 的互异本征值为 $\lambda_1, \lambda_2, \cdots, \lambda_m$，对应的代数重数为 $d_1, d_2, \cdots, d_m$.
T 的特征多项式 $p(\lambda)=\prod_{i=1}^{m}(\lambda-\lambda_i)^{d_i}$ 是零化多项式 (Cayley-Hamilton)
对应不变子空间直和分解 $V=Null\,p(T)=\bigoplus_{i=1}^{s} Null(T-\lambda_i I)^{d_i}$.
记 λ_i 对应的广义本征空间 $G(\lambda_i, T)=Null(T-\lambda_i I)^{d_i}$.
在每个 $G(\lambda_i, T)$ 中，$(T-\lambda_i I)|_{G(\lambda_i,T)}$ 是幂零算子.
（显然，$[(T-\lambda_i I)|_{G(\lambda_i,T)}]^{d_i}=0$）
由 5-42，算子 $T|_{G(\lambda_i,T)}$ 在 $G(\lambda_i, T)$ 的 Jordan 基下成 Jordan 形.
又由 $V=\bigoplus_{i=1}^{s} G(\lambda_i, T)$，把各 $G(\lambda_i, T)$ 的 Jordan 基并成 V 的基，
T 在该基下成 Jordan 形矩阵. Q.E.D.

- 设 T 是有限维复向量空间 V 上的算子，T 在 Jordan 基下成 Jordan 形，且本征值 λ_i 对应的 t 阶 Jordan 块的个数

$$n_i(t)=rank(T-\lambda_i I)^{t-1}+rank(T-\lambda_i I)^{t+1}-2rank(T-\lambda_i I)^{t} \quad (5\text{-}44)$$

证：$\forall m\in \mathbb{N}^*$，$Null(T-\lambda_i I)^m \subset G(\lambda_i, T)$
$\Rightarrow Null(T-\lambda_i I)^m = Null(T-\lambda_i I)^m|_{G(\lambda_i,T)} = Null[(T-\lambda_i I)|_{G(\lambda_i,T)}]^m$
由 5-43，$(T-\lambda_i I)|_{G(\lambda_i,T)}$ 是幂零算子，
由 5-40，$n_i(t)=rank[(T-\lambda_i I)|_{G(\lambda_i,T)}]^{t-1}+rank[(T-\lambda_i I)|_{G(\lambda_i,T)}]^{t+1}$
$\qquad -2rank[(T-\lambda_i I)|_{G(\lambda_i,T)}]^{2t}$
代入线性映射基本定理：

$$n_i(t) = \{\dim G(\lambda_i,T) - \dim Null[(T-\lambda_i I)|_{G(\lambda_i,T)}]^{t-1}\}$$
$$+\{\dim G(\lambda_i,T) - \dim Null[(T-\lambda_i I)|_{G(\lambda_i,T)}]^{t+1}\}$$
$$-2\{\dim G(\lambda_i,T) - \dim Null[(T-\lambda_i I)|_{G(\lambda_i,T)}]^{t}\}$$
$$= 2\dim Null(T-\lambda_i I)^t - \dim Null(T-\lambda_i I)^{t-1} - \dim Null(T-\lambda_i I)^{t+1}$$
$$= [\dim V - \dim Null(T-\lambda_i I)^{t-1}] + [\dim V - \dim Null(T-\lambda_i I)^{t+1}]$$
$$-2[\dim V - \dim Null(T-\lambda_i I)^t]$$
$$= rank(T-\lambda_i I)^{t-1} + rank(T-\lambda_i I)^{t+1} - 2rank(T-\lambda_i I)^t \quad \mathcal{Q.E.D.}$$

※ 推论：若不考虑各 Jordan 块的排列顺序，则算子的 Jordan 形矩阵是唯一的． (5-45)

- $\forall A \in C^{n,n}$，存在唯一的 Jordan 形矩阵 J (再次强调不考虑各 Jordan 块的排列顺序) 与 A 相似，J 称为 A 的 Jordan 标准形．

▸ 求 Jordan 形的算法

- 以上在 $\mathbb{C}$ 中求出标准形的过程，可以推广到任意数域 $\mathbb{F}$，只要求算子特征多项式的所有根都属于 $\mathbb{F}$．
- 设 T 是有限维向量空间 V 中的算子，若存在 T 的 Jordan 形，其算法：

1) 求出 T 在 V 的任一组基 $\vec{b}_1, \vec{b}_2, \cdots, \vec{b}_n$ 下的矩阵 A．

2) 计算特征多项式 $p(\lambda) = \det(\lambda I - A)$

3) 求出 $p(\lambda)$ 的互异根 $\lambda_1, \lambda_2, \cdots, \lambda_m$ (全部属于 $\mathbb{F}$)

4) 对每个 λ_i，由公式

$$n_i(t) = rank(A-\lambda_i I)^{t-1} + rank(A-\lambda_i I)^{t+1} - 2rank(A-\lambda_i I)^t$$

算出本征值 λ_i 对应的 t 阶 Jordan 块的个数．

5) 把 Jordan 块随意排列成准对角形 J，即所求 Jordan 形矩阵．

例 5-4：求 $T(x,y,z) = (2x+6y-15z,\ x+y-5z,\ x+2y-6z)$ 的 Jordan 形．

解：在标准正交基下 T 的矩阵：

$$A=\begin{pmatrix}2 & 6 & -15\\ 1 & 1 & -5\\ 1 & 2 & -6\end{pmatrix}$$

矩阵A的特征多项式：

$$\det(\lambda I-A)=\begin{vmatrix}\lambda-2 & -6 & 15\\ -1 & \lambda-1 & 5\\ -1 & -2 & \lambda+6\end{vmatrix}=(\lambda+1)^3$$

$\Rightarrow$ A有唯一本征值 $\lambda=-1$

对 $\lambda=-1$，

$$A-\lambda I=\begin{pmatrix}3 & 6 & -15\\ 1 & 2 & -5\\ 1 & 2 & -5\end{pmatrix},\ \operatorname{rank}(A-\lambda I)=1$$

$(A-\lambda I)^2=0$

1阶Jordan块 $n(1)=\operatorname{rank}(I)+\operatorname{rank}(A-\lambda I)^2-2\operatorname{rank}(A-\lambda I)=1$

2阶Jordan块 $n(2)=\operatorname{rank}(A-\lambda I)+\operatorname{rank}(A-\lambda I)^3-2\operatorname{rank}(A-\lambda I)^2=1$

显然不存在3阶或以上的Jordan块

于是T的Jordan标准形：

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

- 设J是复向量空间上算子T的Jordan形，本征值 λ_i 对应的Jordan块的阶数之和等于本征值 λ_i 的代数重数 (5.46)

证：T的特征多项式 $p(\lambda)=\det(\lambda I-J)=\prod_{i=1}^{m}(\lambda-\lambda_i)^{n_i}$

又 $p(\lambda)=\prod_{i=1}^{m}(\lambda-\lambda_i)^{d_i}$ $\Rightarrow n_i=d_i$ Q.E.D.

※ 推论：λ_i 对应的广义本征空间维数 $\dim G(\lambda_i,T)$ 就是 λ_i 的代数重数.

- Jordan-Chevalley 分解定理
- 设 T 是复向量空间 V 上的算子，V 中存在唯一的可对角化算子 D 与幂零算子 N，使得 $T=D+N$，且 $DN=ND$． (5-47)

证：设 T 的互异本征值为 $\lambda_1,\cdots,\lambda_m$，对应的代数重数为 $d_1,\cdots,d_m$．

分解的存在性是显然的：

$$V=\bigoplus_{i=1}^{m}\mathrm{Null}(T-\lambda_i I)^{d_i}=\bigoplus_{i=1}^{m}G(\lambda_i,T)$$

令 D 是 V 上算子，使得 $D|_{G(\lambda_i,T)}=\lambda_i I$

D 在 T 的 Jordan 基下成对角形；

$N|_{G(\lambda_i,T)}=T|_{G(\lambda_i,T)}-D|_{G(\lambda_i,T)}=(T-\lambda_i I)|_{G(\lambda_i,T)}$ 是 $G(\lambda_i,T)$ 上幂零算子

设其幂零指数为 l_i．

那么 $N^{\max\{l_1,\cdots,l_m\}}=0$，$N$ 是 V 上幂零算子．

$D|_{G(\lambda_i,T)}=\lambda_i T^0|_{G(\lambda_i,T)}$，$N|_{G(\lambda_i,T)}=(T^1-\lambda_i T^0)|_{G(\lambda_i,T)}$

$\Rightarrow D|_{G(\lambda_i,T)}$ 与 $N|_{G(\lambda_i,T)}$ 可交换 $\Rightarrow$ D 与 N 可交换．

分解的唯一性：

设 V 上又有可对角化算子 D_1 与幂零算子 N_1 使 $T=D_1+N_1$ 且 $D_1N_1=N_1D_1$．

$TD=(D+N)D=D(D+N)=DT\Rightarrow T$ 与 D 可交换．

同理有 T 与 N，D_1，N_1 都可交换．

$\forall\vec{v}\in G(\lambda_i,T)$，$(A-\lambda_i I)^{d_i}D_1(\vec{v})=D_1(A-\lambda_i I)^{d_i}(\vec{v})=0$．

$$(A-\lambda_i I)^{d_i}N_1(\vec{v})=N_1(A-\lambda_i I)^{d_i}(\vec{v})=0$$

$\Rightarrow G(\lambda_i,T)$ 是 A，D，N，D_1，N_1 的公共不变子空间．

$N|_{G(\lambda_i,T)}$ 与 $N_1|_{G(\lambda_i,T)}$ 幂零，则 $(N-N_1)|_{G(\lambda_i,T)}$ 也幂零．

$\Rightarrow D_1|_{G(\lambda_i,T)}=(D+N-N_1)|_{G(\lambda_i,T)}=\lambda_i I+(N-N_1)|_{G(\lambda_i,T)}$．

则 $G(\lambda_i,T)$ 存有一组基使 D_1 成 Jordan 形，其对角元全为 λ_i．

$\Rightarrow D_1|_{G(\lambda_i,T)}$ 的本征值为 λ_i． $\Rightarrow D_1|_{G(\lambda_i,T)}=\lambda_i I=D|_{G(\lambda_i,T)}$

$\Rightarrow D=D_1$，$N=N_1$，即证明分解的唯一性． Q.E.D.

※ Jordan-Chevalley 分解的应用.

Taylor 展开求算子的函数：$f(x+a)=\sum_{k=0}^{\infty}\frac{f^{(k)}(a)}{k!}x^k$

形式化地用算子 T 代替变元 $x+a$.

$\Rightarrow f(T)=f(D+N)=\sum_{k=0}^{\infty}\frac{f^{(k)}(D)}{k!}N^k=\sum_{k=0}^{l-1}\frac{f^{(k)}(D)}{k!}N^k$ (5-48)

其中 D 是可对角化算子，N 是幂零指数为 l 的幂零算子.

若 $f(x)=e^x$，有：$e^T=e^D\sum_{k=0}^{l-1}\frac{1}{k!}N^k$

二项式定理求算子的幂：

$T^n=(D+N)^n=\sum_{k=0}^{l-1}C_n^kD^{n-k}N^k$. (5-49)

5.7 广义本征空间

- 本节内容与 §5.4 完全平行. 即完全从广义本征空间的角度出发，得到不变子空间的直和分解的结论. 为此要先准确定义算子的广义本征空间.

▸ 零空间序列

- 设 $T\in\mathcal{L}(V)$，则 $Null\,T^k\subset Null\,T^{k+1}$，$\forall k=0,1,2,\cdots$ (5-50)

证：$\forall\vec{v}\in Null\,T^k$，$T^k(\vec{v})=0$

$\Rightarrow T^{k+1}(\vec{v})=T[T^k(\vec{v})]=0\Rightarrow Null\,T^k\subset Null\,T^{k+1}$ Q.E.D.

- 设 $T\in\mathcal{L}(V)$，若 $\exists k\in\mathbb{N}^*$ s.t. $Null\,T^k=Null\,T^{k+1}$，则有 $\forall r\in\mathbb{N}^*$，$Null\,T^k=Null\,T^{k+r}$ (5-51)

证：$\forall r\in\mathbb{N}^*$，已有 $Null\,T^{k+r}\subset Null\,T^{k+r+1}$.

$\forall\vec{v}\in Null\,T^{k+r+1}$，$T^{k+r+1}(\vec{v})=0\Rightarrow T^{k+1}[T^r(\vec{v})]=0$

$\Rightarrow T^r(\vec{v})\in Null\,T^{k+1}=Null\,T^k\Rightarrow T^k[T^r(\vec{v})]=T^{k+r}(\vec{v})=0$

$\Rightarrow\vec{v}\in Null\,T^{k+r}\Rightarrow Null\,T^{k+r+1}\subset Null\,T^{k+r}$

$\Rightarrow Null\,T^{k+r}=Null\,T^{k+r+1}$

归纳可得 $\forall r\in\mathbb{N}^*$，$Null\,T^k=Null\,T^{k+r}$ Q.E.D.

- 设 $T\in\mathcal{L}(V)$，$\forall r\in\mathbb{N}^*$，$Null\,T^{\dim V}=Null\,T^{\dim V+r}$. (5-52)

证：假设 $Null\,T^{\dim V} \neq Null\,T^{\dim V+1}$

则有 $\{0\} = Null\,T^0 \subsetneq Null\,T^1 \subsetneq \cdots \subsetneq Null\,T^{\dim V} \subsetneq Null\,T^{\dim V+1}$

$\Rightarrow 0 = \dim Null\,T^0 < \dim Null\,T^1 < \cdots < \dim Null\,T^{\dim V+1}$

$\Rightarrow \dim Null\,T^{\dim V+1} \geq \dim V + 1$，这是不可能的

$\Rightarrow Null\,T^{\dim V} = Null\,T^{\dim V+1}$，由5-51，定理得证. Q.E.D.

- 广义本征空间
  - $T \in \mathcal{L}(V)$，设 λ_i 是T的本征值，$\vec{v}$ 称为T的对应 λ_i 的广义本征向量，若 $\exists k \in \mathbb{N}^+$ s.t. $(T-\lambda_i I)^k(\vec{v}) = 0$.
  - 所有对应 λ_i 的广义本征向量的集合称为T的对应 λ_i 的广义本征空间，记为 $G(\lambda_i, T)$.

    （一些国内教材称之为"根子空间"）.
  - 设 $T \in \mathcal{L}(V)$，则 $G(\lambda_i, T) = Null(T-\lambda_i I)^{\dim V}$ (5-52)

证：$\exists k \in \mathbb{N}^+$ s.t. $(T-\lambda_i I)^k(\vec{v}) = 0 \Rightarrow \vec{v} \in Null(T-\lambda_i I)^k \subset Null(T-\lambda_i I)^{\dim V}$

$\Rightarrow G(\lambda_i, T) = \bigcup_{k \in \mathbb{N}^+} Null(T-\lambda_i I)^k = Null(T-\lambda_i I)^{\dim V}$ Q.E.D.

※ 注意到§5.4中我们定义广义本征空间 $G'(\lambda_i, T) = Null(T-\lambda_i I)^{d_i}$，其中 d_i 是 λ_i 对应的重数，$d_i \leq \dim V$，则有 $G'(\lambda_i, T) \subset G(\lambda_i, T)$. 最终我们的目标是证明两种定义等价，即 $G'(\lambda_i, T) = G(\lambda_i, T)$.

- 属于不同本征值的广义本征向量线性无关（5-1的推广）

  设 $T \in \mathcal{L}(V)$，$\lambda_1, \lambda_2, \cdots, \lambda_m$ 是T的互异本征值，$\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_m$ 是对应的广义本征向量，则 $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_m$ 线性无关. (5-53)

证：设 $n = \dim V$，设 $k \in \mathbb{N}^*$ s.t. $(T-\lambda_1 I)^k(\vec{v}_1) \neq 0$ & $(T-\lambda_1 I)^{k+1}(\vec{v}_1) = 0$.

令 $\vec{w}_1 = (T-\lambda_1 I)^k(\vec{v}_1) \neq 0 \Rightarrow (T-\lambda_1 I)(\vec{w}_1) = 0 \Rightarrow T(\vec{w}_1) = \lambda_1 \vec{w}_1$.

容易推广得 $p(T)(\vec{w}_1) = p(\lambda_1)\vec{w}_1$

$\exists a_1, a_2, \cdots, a_m \in \mathbb{F}$ s.t. $\sum_{i=1}^m a_i \vec{v}_i = 0$

$\Rightarrow 0 = \sum_{i=1}^m a_i (T-\lambda_1 I)^k (T-\lambda_2 I)^n (T-\lambda_3 I)^n \cdots (T-\lambda_m I)^n (\vec{v}_i)$

$\forall i=2,3,\cdots,n$ ，$\vec{v_i}\in G(\lambda_i,T)\Rightarrow(T-\lambda_i I)^n(\vec{v_i})=0$

$\Rightarrow 0=a_1(T-\lambda_1 I)^k(T-\lambda_2 I)^n\cdots(T-\lambda_m I)^n(\vec{v_1})$

$=a_1(T-\lambda_2 I)^n\cdots(T-\lambda_m I)^n(\vec{w_1})$

$=a_1(\lambda_1-\lambda_2)^n\cdots(\lambda_1-\lambda_m)^n\vec{w_1}$

$\Rightarrow a_1=0$ ，同理有 $a_2=\cdots=a_m=0$

即 $\vec{v_1},\vec{v_2},\cdots,\vec{v_m}$ 线性无关． 2.5.Q

- $\mathbb{C}$上算子的广义本征空间的直和分解．

· 设T是复向量空间V上算子，$\lambda_1,\lambda_2,\cdots,\lambda_m$是T的互异本征值． (5-54)

i) $V=\bigoplus_{i=1}^{m}G(\lambda_i,T)$

ii) 每个$G(\lambda_i,T)$都是T的不变子空间

iii) 每个$(T-\lambda_i I)|_{G(\lambda_i,T)}$都是幂零算子

证：ii) 与 iii) 显然 (见5-23 & 5-43)

i) 对 dim V 作归纳法：

dim V = 1 时显然；

假设 dim V < n 时有 $V=\bigoplus_{i=1}^{m}G(\lambda_i,T)$ ；

当 dim V = n 时：

T有本征值λ_1，$G(\lambda_1,T)$是T的不变子空间．

下面证明 $V=Null(T-\lambda_1 I)^n\oplus Ran(T-\lambda_1 I)^n$ ：

$\forall\vec{v}\in Null(T-\lambda_1 I)^n\oplus Ran(T-\lambda_1 I)^n$，

$(T-\lambda_1 I)^n(\vec{v})=0$ 且 $\exists\vec{u}\in V$ s.t. $(T-\lambda_1 I)^n(\vec{u})=\vec{v}$

$\Rightarrow(T-\lambda_1 I)^{2n}(\vec{u})=0\Rightarrow\vec{u}\in Null(T-\lambda_1 I)^{2n}=Null(T-\lambda_1 I)^n$

$\Rightarrow(T-\lambda_1 I)^n(\vec{u})=0\Rightarrow\vec{v}=0$

$\Rightarrow Null(T-\lambda_1 I)^n\cap Ran(T-\lambda_1 I)^n=\{0\}$

由线性映射基本定理：$\dim(Null(T-\lambda_1 I)^n\oplus Ran(T-\lambda_1 I)^n)$

$=\dim Null(T-\lambda_1 I)^n+\dim Ran(T-\lambda_1 I)^n=\dim V$

于是 $V = G(\lambda_1, T) \oplus U$，其中 $U = \mathrm{Ran}(T-\lambda_1 I)^n$

易证 U 是 T 的不变子空间：

$\forall \vec{v} \in \mathrm{Ran}(T-\lambda_1 I)^n$，$\exists \vec{u} \in V$ s.t. $(T-\lambda_1 I)^n(\vec{u}) = \vec{v}$

$T(\vec{v}) = T(T-\lambda_1 I)^n(\vec{u}) = (T-\lambda_1 I)^n[T(\vec{u})] \in \mathrm{Ran}(T-\lambda_1 I)^n$.

$\dim U < n$，因此可以对 $T|_U$ 应用归纳假设：

T 所有对应 λ_1 的广义本征向量都在 $G(\lambda_1, T)$ 中，那么 $T|_U$ 的本征值为 $\lambda_2, \lambda_3, \cdots, \lambda_m$

$\Rightarrow U = \bigoplus_{i=2}^{m} G(\lambda_i, T|_U)$.

下面证明 $\forall i = 2, 3, \cdots, m$，$G(\lambda_i, T) = G(\lambda_i, T|_U)$：

$G(\lambda_i, T|_U) \subset G(\lambda_i, T)$ 显然；

由 $V = G(\lambda_1, T) \oplus G(\lambda_2, T|_U) \oplus \cdots \oplus G(\lambda_m, T|_U)$ 得，

$\forall \vec{v} \in G(\lambda_i, T)$，$\vec{v} = \vec{v}_1 + \vec{v}_2 + \cdots + \vec{v}_m$

其中 $\vec{v}_1 \in G(\lambda_1, T)$，$\vec{v}_j \in G(\lambda_j, T|_U)$，$j = 2, \cdots, m$.

由 5-53，$\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_m$ 线性无关.

$\Rightarrow \vec{v} = \vec{v}_i \in G(\lambda_i, T|_U) \Rightarrow G(\lambda_i, T) \subset G(\lambda_i, T|_U)$

$\Rightarrow G(\lambda_i, T) = G(\lambda_i, T|_U)$.

$\Rightarrow U = \bigoplus_{i=2}^{m} G(\lambda_i, T)$，$V = G(\lambda_1, T) \oplus U = \bigoplus_{i=1}^{m} G(\lambda_i, T)$ Q.E.D.

※ 算子的本征向量不一定能构成空间的基。算子可对角化的充要条件就是空间中有一组本征向量构成的基。然而，广义本征向量总是足够构成基。

· 设 T 是复向量空间 V 上算子，则 V 有一组由 T 的广义本征向量构成的基.

证：$V = \bigoplus_{i=1}^{m} G(\lambda_i, T)$，在每个 $G(\lambda_i, T)$ 中取一组基，并成 V 的基. (5-55)

※ 由 §5.5，$(T-\lambda_i I)|_{G(\lambda_i, T)}$ 是幂零算子，则 $G(\lambda_i, T)$ 上有一组 Jordan 基使 $(T-\lambda_i I)|_{G(\lambda_i, T)}$ 成 Jordan 形；把每个 $G(\lambda_i, T)$ 的 Jordan 基并成 V 的基，T 在该基下成 Jordan 形矩阵. (5-43)

▸ 代数重数.

- 设T是复向量空间V上算子，$\lambda_1, \lambda_2, \cdots, \lambda_m$是T的互异本征值，$d_1, d_2, \cdots, d_m$是对应的代数重数，则 $\dim G(\lambda_i, T) = d_i$ (5-56)

证：由5-54，$V = \bigoplus_{i=1}^{m} G(\lambda_i, T)$.

由5-20，T有准对角形 $diag(A_1, A_2, \cdots, A_m)$.

对应 $T|_{G(\lambda_i, T)}$ 的矩阵 A_i.

A_i 有唯一本征值 $\lambda_i \Rightarrow \det(\lambda I - A_i) = (\lambda - \lambda_i)^{\dim G(\lambda_i, T)}$

$\Rightarrow \det(\lambda I - T) = \det(\lambda I - A) = \det(\lambda I - A_1) \cdots \det(\lambda I - A_m)$

$= (\lambda - \lambda_1)^{\dim G(\lambda_1, T)} \cdots (\lambda - \lambda_m)^{\dim G(\lambda_m, T)}$

又由§5.1结论 $\det(\lambda I - T) = (\lambda - \lambda_1)^{d_1} \cdots (\lambda - \lambda_m)^{d_m}$

$\Rightarrow \dim G(\lambda_i, T) = d_i$. Q.E.D.

※ 这个结论本质上就是5-46.

同时说明 $G(\lambda_i, T) = Null(T - \lambda_i I)^{d_i} = Null(T - \lambda_i I)^{\dim V}$.

两种广义本征空间的定义是等价的.

- 本征值的重数

λ_i的代数重数 $= \dim G(\lambda_i, T) = \dim Null(T - \lambda_i I)^{\dim V}$

λ_i的几何重数 $= \dim E(\lambda_i, T) = \dim Null(T - \lambda_i I)$

几何重数 ≤ 代数重数， $E(\lambda_i, T) \subset G(\lambda_i, T)$

算子可对角化 $\Longleftrightarrow$ 几何重数 = 代数重数

$\Longleftrightarrow E(\lambda_i, T) = G(\lambda_i, T)$

$\Longleftrightarrow V = \bigoplus_{i=1}^{m} E(\lambda_i, T)$

5.8 复化

- 利用实向量空间V构造复向量空间V_C，从而把一些复向量空间成立的定理推广到实向量空间.

▸ 复化

- 设 V 是实向量空间，V 的复化 $V_{\mathbb{C}} = V \times V$，其元素为 $(\vec{u}, \vec{v})$，$\vec{u}, \vec{v} \in V$.

  定义 $V_{\mathbb{C}}$ 上加法：$(\vec{u_1}, \vec{v_1}) + (\vec{u_2}, \vec{v_2}) = (\vec{u_1} + \vec{u_2}, \vec{v_1} + \vec{v_2})$.

  $V_{\mathbb{C}}$ 上与 i 的乘法：$i(\vec{u}, \vec{v}) = (-\vec{v}, \vec{u})$.

  利用结合律，$V_{\mathbb{C}}$ 上数乘：$(a+bi)(\vec{u}, \vec{v}) = (a\vec{u} - b\vec{v}, a\vec{v} + b\vec{u})$，$a, b \in \mathbb{R}$.

※ 如果把 $V_{\mathbb{C}}$ 仍看作实向量空间，与复数的乘法等价于作用矩阵 $\mathbb{R}^{2,2}$：

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$ （恰好是 $\mathbb{R}^2$ 上正规算子的矩阵，why？）

- 把 $(\vec{u}, \vec{v})$ 写作 $\vec{u} + i\vec{v}$，则 $V_{\mathbb{C}}$ 关于以上定义的加法和数乘成复向量空间.

- 设 V 是实向量空间，$V_{\mathbb{C}}$ 是 V 的复化（作为复向量空间）. (5-57)

  i) 若 $\vec{v_1}, \vec{v_2}, \cdots, \vec{v_n}$ 是 V 的基，则它也是 $V_{\mathbb{C}}$ 的基.

  ii) $\dim V = \dim V_{\mathbb{C}}$

证：i) 若 $V = \mathrm{span}(\vec{v_1}, \vec{v_2}, \cdots, \vec{v_n})$，则 $\forall \vec{u} + i\vec{v} \in V_{\mathbb{C}}$，$\vec{u}, \vec{v} \in \mathrm{span}(\vec{v_1}, \cdots, \vec{v_n})$

$\Rightarrow V_{\mathbb{C}} = \mathrm{span}(\vec{v_1}, \vec{v_2}, \cdots, \vec{v_n})$.

$\exists \lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{C}$ s.t. $\sum_{i=1}^{n} \lambda_i \vec{v_i} = 0$.

$\Rightarrow \sum_{i=1}^{n} (\mathrm{Re}\lambda_i)\vec{v_i} = 0$ & $\sum_{i=1}^{n} (\mathrm{Im}\lambda_i)\vec{v_i} = 0$.

$\vec{v_1}, \cdots, \vec{v_n}$ 在 V 上线性无关 $\Rightarrow \mathrm{Re}\lambda_1 = \cdots = \mathrm{Re}\lambda_n = 0$，$\mathrm{Im}\lambda_1 = \cdots = \mathrm{Im}\lambda_n = 0$

$\Rightarrow \vec{v_1}, \cdots, \vec{v_n}$ 在 $V_{\mathbb{C}}$ 上线性无关.

即 $\vec{v_1}, \cdots, \vec{v_n}$ 构成 $V_{\mathbb{C}}$ 的一组基.

ii) 由 i) 立证. Q.E.D.

▸ 算子的复化

- 设 V 是实向量空间，$T \in \mathcal{L}(V)$，T 的复化 $T_{\mathbb{C}} \in \mathcal{L}(V_{\mathbb{C}})$ 定义为：

  $T_{\mathbb{C}}(\vec{u} + i\vec{v}) = T(\vec{u}) + iT(\vec{v})$，$\forall \vec{u}, \vec{v} \in V$. (5-58)

验证 $T_{\mathbb{C}} \in \mathcal{L}(V_{\mathbb{C}})$，重点是复数乘法：

$\forall a, b \in \mathbb{R}$，$T_{\mathbb{C}}[(a+bi)(\vec{u} + i\vec{v})] = T_{\mathbb{C}}[(a\vec{u} - b\vec{v}) + i(a\vec{v} + b\vec{u})]$

$= T(a\vec{u} - b\vec{v}) + iT(a\vec{v} + b\vec{u}) = (a+bi)T(\vec{u}) + (ai - b)T(\vec{v})$

$= (a+bi)[T(\vec{u})+iT(\vec{v})] = (a+bi)T_{\mathbb{C}}(\vec{u}+i\vec{v})$. Q.E.D.

- 设 $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n$ 是实向量空间 V 的基（也是 $V_{\mathbb{C}}$ 的基），则 T 与 $T_{\mathbb{C}}$ 关于该基有相同的矩阵. (5-59)

※ 算子的复化也可以用矩阵定义：$T\in\mathcal{L}(V)$ 在某组基下有矩阵 $A\in\mathbb{R}^{n,n}$，若将 A 看作 $\mathbb{C}^{n,n}$ 的元素，则 A 对应 $T_{\mathbb{C}}\in\mathcal{L}(V_{\mathbb{C}})$ 在该基下的矩阵. (5-60)

这个操作之前已经多次使用过了.

- T 的行列式等于 $T_{\mathbb{C}}$ 的行列式. (5-61)
- T 的特征多项式等于 $T_{\mathbb{C}}$ 的特征多项式. (5-62)
- （实向量空间的）Cayley-Hamilton 定理.

$\mathbb{R}$ 上算子 T 的特征多项式是零化多项式. (5-63)

证：复化 $T_{\mathbb{C}}$ 是 $\mathbb{C}$ 上算子，由 5-26，$p(T_{\mathbb{C}})=0 \Rightarrow p(T)=0$. Q.E.D.

- 设 T 是实向量空间 V 上算子，T 有一维或二维不变子空间. (5-64)

证：T 的复化 $T_{\mathbb{C}}$ 有本征值 $a+bi$，$a,b\in\mathbb{R}$.

$\Rightarrow \exists \vec{u},\vec{v}\in V-\{0\}$, s.t. $T_{\mathbb{C}}(\vec{u}+i\vec{v}) = (a+bi)(\vec{u}+i\vec{v})$

$\Rightarrow T(\vec{u})+iT(\vec{v}) = (a\vec{u}-b\vec{v})+(a\vec{v}+b\vec{u})i$

$\Rightarrow T(\vec{u}) = a\vec{u}-b\vec{v}$, $T(\vec{v}) = (a\vec{v}+b\vec{u})$

可见 $\mathrm{span}(\vec{u},\vec{v})$ 是 T 的不变子空间，且 $0<\dim\mathrm{span}(\vec{u},\vec{v})\leq 2$. Q.E.D.

▸ 复化的本征值

- $T_{\mathbb{C}}$ 的实本征值等于 T 的本征值. (5-65)

证：i) 若 $\lambda\in\mathbb{R}$ 是 $T\in\mathcal{L}(V)$ 的本征值，$\exists \vec{v}\in V-\{0\}$ s.t. $T(\vec{v})=\lambda\vec{v}$

$\Rightarrow T_{\mathbb{C}}(\vec{v})=\lambda\vec{v} \Rightarrow \lambda$ 是 $T_{\mathbb{C}}\in\mathcal{L}(V_{\mathbb{C}})$ 的本征值.

ii) 若 $\lambda\in\mathbb{R}$ 是 $T_{\mathbb{C}}\in\mathcal{L}(V_{\mathbb{C}})$ 的本征值，$\exists \vec{u},\vec{v}\in V-\{0\}$ s.t. $T_{\mathbb{C}}(\vec{u}+i\vec{v})=\lambda(\vec{u}+i\vec{v})$

$\Rightarrow T(\vec{u})=\lambda\vec{u}$, $T(\vec{v})=\lambda\vec{v} \Rightarrow \lambda$ 是 $T\in\mathcal{L}(V)$ 的本征值. Q.E.D.

（上面都是废话，从特征多项式看很显然）

- $T_{\mathbb{C}}$ 的复本征值两两共轭.

$\lambda \in \mathbb{C}$ 是 $T_{\mathbb{C}} \in \mathcal{L}(V_{\mathbb{C}})$ 的本征值，则 $\bar{\lambda}$ 也是 $T_{\mathbb{C}}$ 的本征值．且 λ 的重数等于 $\bar{\lambda}$ 的重数． (5-66)

证：$T_{\mathbb{C}}$ 的特征多项式是实系数多项式，其复根两两共轭． Q.E.D.

• 奇数维实向量空间上的算子必有本征值． (5-67)

证：设 T 是奇数维实向量空间 V 上算子，$T_{\mathbb{C}}$ 的虚本征值总是成对出现且重数相等，则所有虚本征值重数之和为偶数．即 $T_{\mathbb{C}}$ 有实本征值，对应 T 有本征值．

Q.E.D.

双线性形式 & 二次型

6.1 双线性形式

- 设 V 是 $\mathbb{F}$ 上向量空间，映射 $f: V\times V\to F$ 称为双线性形式，若满足：

  i) $\forall \lambda_1,\lambda_2\in\mathbb{F}$，$\vec{v_1},\vec{v_2},\vec{w}\in V$，

  $$f(\lambda_1\vec{v_1}+\lambda_2\vec{v_2},\vec{w})=\lambda_1 f(\vec{v_1},\vec{w})+\lambda_2 f(\vec{v_2},\vec{w})$$

  ii) $\forall \xi_1,\xi_2\in\mathbb{F}$，$\vec{v},\vec{w_1},\vec{w_2}\in V$，

  $$f(\vec{v},\xi_1\vec{w_1}+\xi_2\vec{w_2})=\xi_1 f(\vec{v},\vec{w_1})+\xi_2 f(\vec{v},\vec{w_2})$$

- 设 V 有一组基 $\vec{b_1},\vec{b_2},\cdots,\vec{b_n}$，将 $f(\vec{v},\vec{w})$ 按基底展开：

  $$f(\vec{v},\vec{w})=f\left(\sum_{i=1}^{n}x_i\vec{b_i},\sum_{j=1}^{n}y_j\vec{b_j}\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}x_iy_jf(\vec{b_i},\vec{b_j})$$

  即 $f(\vec{v},\vec{w})$ 由它在基上的像唯一确定.

- 双线性形式构成的向量空间 $\mathcal{L}(V\times V,\mathbb{F})$ 与矩阵空间 $\mathbb{F}^{n,n}$ 同构.

  即任给 $A=\{a_{ij}\}_{i,j=1}^{n}\in\mathbb{F}^{n,n}$ 都有 $f(\vec{v},\vec{w})\in\mathcal{L}(V\times V,F)$ 与之一一对应.

  取 V 的基 $\vec{b_1},\vec{b_2},\cdots,\vec{b_n}$，有：$f(\vec{b_i},\vec{b_j})=a_{i,j}$.

  $$\Rightarrow f(\vec{v},\vec{w})=\sum_{i=1}^{n}\sum_{j=1}^{n}x_iy_ia_{i,j}$$

- $$A=\begin{pmatrix} f(\vec{b_1},\vec{b_1}) & f(\vec{b_1},\vec{b_2}) & \cdots & f(\vec{b_1},\vec{b_n})\\ f(\vec{b_2},\vec{b_1}) & f(\vec{b_2},\vec{b_2}) & \cdots & f(\vec{b_2},\vec{b_n})\\ \vdots & \vdots & \ddots & \vdots\\ f(\vec{b_n},\vec{b_1}) & f(\vec{b_n},\vec{b_2}) & \cdots & f(\vec{b_n},\vec{b_n})\end{pmatrix}$$

  称为双线性形 $f(\vec{v},\vec{w})$ 在基 $\vec{b_1},\vec{b_2},\cdots,\vec{b_n}$ 下的矩阵.

- 引入 $\vec{v}$ 的坐标向量 $\vec{x}=(x_1\cdots x_n)^T$，$\vec{w}$ 的坐标向量 $\vec{y}=(y_1\cdots y_n)^T$

  $f(\vec{v},\vec{w})=\sum_{i=1}^{n}\sum_{j=1}^{n}x_iy_ja_{i,j}$ 可形式化地记为

  $$f(\vec{v},\vec{w})=\vec{x}^TA\vec{y}$$ （或 $f(\vec{v},\vec{w})=X^TAY$ 好看一点 ...）

- 双线性形的基变换

  设 V 有两组基 $\vec{\varepsilon_1},\vec{\varepsilon_2},\cdots,\vec{\varepsilon_n}$ 与 $\vec{\eta_1},\vec{\eta_2},\cdots,\vec{\eta_n}$，双线性形 $f(\vec{v},\vec{w})$ 在两组基下的矩阵分别为 $A=\{a_{i,j}\}$ 和 $B=\{b_{i,j}\}$.

$$\vec{v} = \sum_{i=1}^{n} x_i \vec{\varepsilon}_i \quad \Rightarrow \vec{x} = (x_1 \cdots x_n)^T$$
$$= \sum_{i=1}^{n} x_i' \vec{\eta}_i \quad \Rightarrow \vec{x'} = (x_1' \cdots x_n')^T$$
$$\vec{w} = \sum_{j=1}^{n} y_j \vec{\varepsilon}_j \quad \Rightarrow \vec{y} = (y_1 \cdots y_n)^T$$
$$= \sum_{j=1}^{n} y_j' \vec{\eta}_j \quad \Rightarrow \vec{y'} = (y_1' \cdots y_n')^T$$

设两组基之间过渡矩阵 P ： $\vec{x} = P\vec{x'}$, $\vec{y} = P\vec{y'}$

$$\Rightarrow f(\vec{v},\vec{w}) = \vec{x}^T A \vec{y} = (P\vec{x'})^T A (P\vec{y'}) = \vec{x'}^T (P^T A P) \vec{y'}$$
$$= \vec{x'}^T B \vec{y'}$$

$\Rightarrow B = P^T A P$ ，称为 A 与 B 合同（P 可逆）

- 矩阵之间的合同关系是一种等价关系，相互合同的矩阵构成合同类.（对比矩阵的相似类： $B = P^{-1}AP$ ，对应不同基下的线性映射的关系）

▶ 双线性形式的秩.

- 双线性形式的秩定义为它在某组基下的矩阵的秩（合同矩阵秩相等）
- 若该矩阵满秩，则对应非退化双线性形式.

6.2 二次型

▶ 对称双线性形式

- 若 $\forall \vec{v},\vec{w} \in V$, $f(\vec{v},\vec{w}) = f(\vec{w},\vec{v})$ ，则称 $f(\vec{v},\vec{w})$ 是对称双线性形式.
- 对对称双线性形式 $f(\vec{v},\vec{w})$ ，定义二次型 $Q_f(\vec{v}) = f(\vec{v},\vec{v})$
- 在 V 内取一组基 $\vec{b}_1, \vec{b}_2, \cdots, \vec{b}_n$ ， $f(\vec{v},\vec{w})$ 在该基下的矩阵 $A = \{a_{ij}\}$
  有 $Q_f(\vec{v}) = f(\vec{v},\vec{v}) = \vec{x}^T A \vec{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j \quad (a_{i,j} = a_{j,i})$
- 对称双线性形式对应的矩阵是对称矩阵.
- $Q_f(\vec{v}+\vec{w}) = f(\vec{v}+\vec{w}, \vec{v}+\vec{w}) = f(\vec{v},\vec{v}) + 2f(\vec{v},\vec{w}) + f(\vec{w},\vec{w})$
  $= Q_f(\vec{v}) + 2f(\vec{v},\vec{w}) + Q_f(\vec{w})$.

$$\Rightarrow f(\vec{v},\vec{w}) = \frac{1}{2}\left[Q_f(\vec{v}+\vec{w}) - Q_f(\vec{v}) - Q_f(\vec{w})\right].$$

即二次型也唯一决定对称双线性形式

- $\mathbb{F}$上的对称矩阵都合同于对角矩阵.（合同对角化） (6-1)

证：对维数作归纳法.

i) 1阶对称矩阵显然可对角化.

ii) 假设$(n-1)$阶对称矩阵可对角化

iii) 对n阶对称矩阵，考虑V上的对称双线性形式与之对应
$(\dim V = n)$

取V的一组基$\vec{v_1}, \vec{v_2}, \cdots, \vec{v_n}$使得$f(\vec{v_1}, \vec{v_1}) = \lambda_1 \neq 0$.

令$\begin{cases} \vec{w_1} = \vec{v_1} \\ \vec{w_i} = \dfrac{f(\vec{v_1}, \vec{v_i})}{\lambda_1}\vec{v_1} - \vec{v_i} \quad (i=2,3,\cdots,n) \end{cases}$ （类似Gram-Schmidt正交化）

易证$\vec{w_1}, \vec{w_2}, \cdots, \vec{w_n}$线性无关，即张成$V$的一组基.

且有$f(\vec{w_1}, \vec{w_i}) = f\left(\vec{v_1}, \dfrac{f(\vec{v_1}, \vec{v_i})}{\lambda_1}\vec{v_1} - \vec{v_i}\right)$

$= \dfrac{f(\vec{v_1}, \vec{v_i})}{\lambda_1} f(\vec{v_1}, \vec{v_1}) - f(\vec{v_1}, \vec{v_i}) = 0 \quad (i=2,3,\cdots,n)$

令$U = \mathrm{span}(\vec{w_2}, \vec{w_3}, \cdots, \vec{w_n})$，$\dim U = n-1$.

由ii)，$f(\vec{v}, \vec{w})$在U上的对称矩阵可对角化

设U上一组基$\vec{u_2}, \vec{u_3}, \cdots, \vec{u_n}$使得$f(\vec{u_i}, \vec{u_j}) = \lambda_i \delta_{ij}$.

$\forall \vec{u} \in U$，$f(\vec{v_1}, \vec{u}) = \sum_{i=2}^{n} k_i f(\vec{v_1}, \vec{w_i}) = 0$.

$\Rightarrow f(\vec{v_1}, \vec{u_i}) = 0$，$(i=2,3,\cdots,n)$

令$\vec{u_1} = \vec{v_1}$，则$\vec{u_1}, \vec{u_2}, \cdots, \vec{u_n}$是$V$的基，且$f(\vec{u_i}, \vec{u_j}) = \lambda_i \delta_{ij}$

即$f(\vec{v}, \vec{w})$在V上的对称矩阵可对角化. Q.E.D.

▶ 数域上的二次型

- $\mathbb{F}$上n个变量$x_1, x_2, \cdots, x_n$的二次型

$$f = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j \quad (a_{ij} = a_{ji})$$

系数矩阵$A = \{a_{ij}\}_{i,j=1}^{n}$是$n$阶对称矩阵，称为二次型的矩阵.

$\mathrm{rank}\, A$称为二次型的秩.

$\Rightarrow f = \vec{x}^T A \vec{x}$

- $\mathbb{F}$上的二次型f就是V上二次型$Q_f(\vec{v})$取定基下的形式.

例6-1 求$f = x_1^2 - 2x_1x_2 + 2x_1x_3 - x_2x_3 + 4x_3^2$对应的矩阵

解:
$$\begin{aligned} f = x_1^2 &\quad - x_1x_2 \quad + \tfrac{3}{2}x_1x_3 \\ - x_2x_1 &\quad\quad\quad\quad - \tfrac{1}{2}x_2x_3 \\ + \tfrac{3}{2}x_3x_1 &\quad - \tfrac{1}{2}x_3x_2 \quad + 4x_3^2 \end{aligned} \Rightarrow A = \begin{pmatrix} 1 & -1 & \frac{3}{2} \\ -1 & 0 & -\frac{1}{2} \\ \frac{3}{2} & -\frac{1}{2} & 4 \end{pmatrix}$$

$$f = (x_1\ x_2\ x_3)\begin{pmatrix} 1 & -1 & \frac{3}{2} \\ -1 & 0 & -\frac{1}{2} \\ \frac{3}{2} & -\frac{1}{2} & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

- 以下集合存在相互同构关系:
  - i) $\mathbb{F}$上全体n阶对称矩阵
  - ii) V上($\dim V = n$)全体对称双线性形式
  - iii) V上($\dim V = n$)全体二次型
  - iv) $\mathbb{F}$上全体n元二次型

▶ 二次型的标准形

- 对$\mathbb{F}$上的二次型f, 用可逆方阵P作坐标变换.

  $\vec{x} = P\vec{y}$

  $f = \vec{x}^T A \vec{x} = (P\vec{y})^T A (P\vec{y}) = \vec{y}^T (P^TAP)\vec{y} = \vec{y}^T B \vec{y}$.

  即$\mathbb{F}$上二次型的坐标变换对应矩阵合同.

- 由(6-1), 每个$\mathbb{F}$上的二次型都能变换为标准形, 其矩阵为对角形.

  $f = d_1z_1^2 + \cdots + d_nz_n^2 = \vec{z}^T D \vec{z}$, $D = \mathrm{diag}(d_1, \cdots, d_n)$

- 计算二次型的标准形的算法

  Step 1: 若二次型中某变量平方项不为零, 则对其配方.

  设$a_{1,1} \neq 0$: $f = a_{11}x_1^2 + 2a_{1,2}x_1x_2 + \cdots + 2a_{1,n}x_1x_n + \sum_{i=2}^{n}\sum_{j=2}^{n} a_{i,j}x_ix_j$.

  $= a_{11}(x_1 + \frac{a_{1,2}}{a_{1,1}}x_2 + \cdots + \frac{a_{1,n}}{a_{1,1}}x_n)^2 + \sum_{i=2}^{n}\sum_{j=2}^{n} b_{i,j}x_ix_j$

换元：
$$\begin{cases} y_1 = x_1 + \frac{a_{12}}{a_{11}} x_2 + \cdots + \frac{a_{1n}}{a_{11}} x_n \\ y_2 = x_2 \\ \cdots\cdots \\ y_n = x_n \end{cases}$$

对应坐标变换

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ & 1 & & 0 \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$\Rightarrow f = a_{11} y_1^2 + \sum_{i=2}^{n}\sum_{j=2}^{n} b_{ij} y_i y_j \quad (b_{ij} = b_{ji})$$

再对$(n-1)$个变量的二次型 $\sum_{i=2}^{n}\sum_{j=2}^{n} b_{ij} y_i y_j$ 重复 Step 1.

Step 2：若所有 $a_{ii} = 0$ 而有一个 $a_{ij} \neq 0 \quad (i<j)$

换元：
$$\begin{cases} x_i = y_i + y_j \\ x_j = y_i - y_j \\ x_k = y_k \quad (k \neq i, j) \end{cases}$$

$a_{ij} x_i x_j = a_{ij} y_i^2 + a_{ij} y_j^2$，回到 Step 1.

重复 Step 1 与 Step 2 直到化为标准形.

※ 每次换元的坐标变换矩阵相乘得总坐标变换矩阵.

例 6-2 化二次型 $f = 2x_1x_2 - 6x_2x_3 + 2x_1x_3$ 为标准形，并求变换矩阵.

解：所有平方项为零而 $a_{12} \neq 0$，换元：

$$\begin{cases} x_1 = y_1 + y_2 \\ x_2 = y_1 - y_2 \\ x_3 = y_3 \end{cases} \qquad P_1 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\Rightarrow f = 2(y_1+y_2)(y_1-y_2) + 2(y_1+y_2)y_3 - 6(y_1-y_2)y_3$$
$$= 2y_1^2 - 2y_2^2 - 4y_1y_3 + 8y_2y_3$$

y_1^2系数不为零，对y_1配方：

$$f=2(y_1-y_3)^2-2y_2^2-2y_3^2+8y_2y_3$$

换元：$\begin{cases} y_1=z_1+z_3 \\ y_2=z_2 \\ y_3=z_3 \end{cases}$ $\quad P_2=\begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}$

$$\Rightarrow f=2z_1^2-2z_2^2+8z_2z_3-2z_3^2$$

对z_2配方：$f=2z_1^2-2(z_2-2z_3)^2+6z_3^2$

换元：$\begin{cases} z_1=w_1 \\ z_2=w_2+2w_3 \\ z_3=w_3 \end{cases}$ $\quad P_3=\begin{pmatrix}1&0&0\\0&1&2\\0&0&1\end{pmatrix}$

$\Rightarrow f=2w_1^2-2w_2^2+6w_3^2$，已化成标准形

总坐标变换矩阵：

$$P=P_1P_2P_3=\begin{pmatrix}1&1&0\\1&-1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&2\\0&0&1\end{pmatrix}=\begin{pmatrix}1&1&3\\1&-1&-1\\0&0&1\end{pmatrix}$$

※ 二次型的标准形不是唯一的.

▶ 合同对角化

- 求二次型的标准形的另一种方法是操作二次型对应的对称矩阵，将其合同对角化.
- 矩阵A经过一次行变换与一次相同的列变换后化为矩阵B，则A与B合同.

证：设行变换对应初等矩阵E，则相同的列变换对应初等矩阵E^T. (6-2)

于是有 $B=EAE^T=(E^T)^TAE^T$，即A与B合同. Q.E.D.

- 合同对角化算法 (6-3)

 1) 写出增广矩阵$(A \,\vdots\, I)$

 2) 对$(A \,\vdots\, I)$行变换，并单独对A进行相同的列变换

3) 重复2)直到A化为对角形D. $(A \vdots I) \sim (D \vdots P)$

此时D即所求对角矩阵，P即坐标变换矩阵.

例6-3 合同对角化: $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

解: 这里不能直接行对换，因为列对换会抵消这个操作.

要先使对角线上出现非零元:

$$(A \vdots I) = \left(\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{array}\right)\begin{matrix} +\frac{1}{2}R_2 \\ \\ \end{matrix} \overset{r}{\sim} \left(\begin{array}{cc|cc} \frac{1}{2} & 1 & 1 & \frac{1}{2} \\ 1 & 0 & 0 & 1 \end{array}\right)$$

$$\overset{c}{\Rightarrow} \left(\begin{array}{cc|cc} 1 & 1 & 1 & \frac{1}{2} \\ 1 & 0 & 0 & 1 \end{array}\right)\begin{matrix} \\ -R_1 \end{matrix} \overset{r}{\sim} \left(\begin{array}{cc|cc} 1 & 1 & 1 & \frac{1}{2} \\ 0 & -1 & -1 & \frac{1}{2} \end{array}\right)$$

$$\overset{c}{\Rightarrow} \left(\begin{array}{cc|cc} 1 & 0 & 1 & \frac{1}{2} \\ 0 & -1 & -1 & \frac{1}{2} \end{array}\right)$$

即 $$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & \frac{1}{2} \\ -1 & \frac{1}{2} \end{pmatrix}$$

6.3 Sylvester惯性定律与正定性

▸ 二次型的规范形

- 设复二次型 $f = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j \ (a_{ij} = a_{ji})$ 有标准形 $f = \sum_{i=1}^{n} d_i z_i^2$.

  作基向量重排，使不为零的 d_i 排在前面.

  即 f 有标准形 $d_1 z_1^2 + \cdots + d_r z_r^2 \quad (r \le n)$

  换元(任意复数都有平方根):

  $$u_i = \begin{cases} \sqrt{d_i}\, z_i, & i = 1, 2, \cdots, r \\ z_i, & i = r+1, \cdots, n. \end{cases}$$

  $\Rightarrow f = u_1^2 + u_2^2 + \cdots + u_r^2$，称为二次型的规范形. (6-4)

- 复二次型的规范形由秩 r 唯一确定; 秩相同的复二次型相互等价.

- 设实二次型 $f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\ (a_{ij}=a_{ji})$ 的秩为 r.
  存在 f 的标准形 $\sum_{i=1}^{r}d_iz_i^2$.
  由于只有非负实数可开方，不妨设 $d_1,d_2,\cdots,d_p>0\ (p\le r)$ 而 $d_{p+1},\cdots,d_r<0$.
  换元(开方)：

$$u_i=\begin{cases}\sqrt{d_i}\,z_i\ , & i=1,2,\cdots,p\\ \sqrt{-d_i}\,z_i\ , & i=p+1,\cdots,r\end{cases}$$

$\Rightarrow f=\sum_{i=1}^{p}u_i^2-\sum_{i=p+1}^{r}u_j^2$，称为二次型的规范形 $(0\le p\le r)$. (6-5)

- Sylvester 惯性定律
  实二次型的规范形是唯一的. (6-6)

证：设实二次型 f 有两个规范形：

$$f=x_1^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_r^2$$
$$=y_1^2+\cdots+y_q^2-y_{q+1}^2-\cdots-y_r^2$$

只需证 $p=q$.

考虑 $\mathbb{R}$ 上向量空间 V 内有基 $\vec{u}_1,\cdots,\vec{u}_n$ 使当 $\vec{v}=\sum_{i=1}^{n}x_i\vec{u}_i$ 时，

$$Q_f(\vec{v})=x_1^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_r^2$$

另一组基 $\vec{w}_1,\cdots,\vec{w}_n$ 使当 $\vec{v}=\sum_{i=1}^{n}y_i\vec{w}_i$ 时，

$$Q_f(\vec{v})=y_1^2+\cdots+y_q^2-y_{q+1}^2-\cdots-y_r^2.$$

$\forall\vec{v}\in\mathrm{span}(\vec{u}_1,\cdots,\vec{u}_p)$，$Q_f(\vec{v})=x_1^2+\cdots+x_p^2\geqslant 0$

$\forall\vec{v}\in\mathrm{span}(\vec{w}_{q+1},\cdots,\vec{w}_n)$，$Q_f(\vec{v})=-y_{q+1}^2-\cdots-y_r^2\le 0$

$\Rightarrow \mathrm{span}(\vec{u}_1,\cdots,\vec{u}_p)\cap\mathrm{span}(\vec{w}_{q+1},\cdots,\vec{w}_n)=\{0\}$.

$\Rightarrow n=\dim V\geqslant\dim(\mathrm{span}(\vec{u}_1,\cdots,\vec{u}_p)+\mathrm{span}(\vec{w}_{q+1},\cdots,\vec{w}_n))$

$=\dim\mathrm{span}(\vec{u}_1,\cdots,\vec{u}_p)+\dim\mathrm{span}(\vec{w}_{q+1},\cdots,\vec{w}_n)$

$=p+(n-q)$

$\Rightarrow p\le q$. 同理有 $q\le p$，即 $p=q$. Q.E.D.

※ p 称为实二次型 f 的正惯性指数，$(r-p)$ 称为负惯性指数.
两者的差 $(2p-r)$ 称为符号差.

- 实二次型的规范形由秩 r 与惯性指数 p 唯一决定. (6-7)

▶ 实二次型的分类

- 实二次型 $Q_f(\vec{v})$ 称为：
  i) 正定的，若 $\forall \vec{v} \neq 0$，$Q_f(\vec{v}) > 0$
  ii) 半正定的，若 $\forall \vec{v} \in V$，$Q_f(\vec{v}) \geq 0$
  iii) 负定的，若 $\forall \vec{v} \neq 0$，$Q_f(\vec{v}) < 0$
  iv) 半负定的，若 $\vec{v} \in V$，$Q_f(\vec{v}) \leq 0$
  v) 不定的，若 $\exists \vec{v_1}, \vec{v_2} \in V$，$Q_f(\vec{v_1}) > 0$，$Q_f(\vec{v_2}) < 0$.

- 设实二次型 $f = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j \ (a_{ij} = a_{ji})$，其对应矩阵为 A，下列表述等价： (6-8)
  i) f 正定 (也说 A 正定)
  ii) $\forall \vec{v} \in V - \{0\}$，$Q_f(\vec{v}) > 0$
  iii) $n =$ 秩 $r =$ 正惯性指数 p
  iv) 在 $\mathbb{R}$ 中，A 合同于 I.

证 由定义有 i) $\Longleftrightarrow$ ii).
iii) $\Rightarrow$ iv)：由 iii) 得 f 有规范形 $u_1^2 + \cdots + u_n^2$，对应矩阵为 I，
则 A 合同于 I.
iv) $\Rightarrow$ iii)：由 iv) 得 f 等价于 $u_1^2 + \cdots + u_n^2$，即 $n = r = p$.
iii) $\Rightarrow$ ii)：由 iii) 得 f 有规范形 $u_1^2 + \cdots + u_n^2$，显然 $Q_f(\vec{v}) > 0 \quad (\vec{v} \neq 0)$
ii) $\Rightarrow$ iii)：设 $Q_f(\vec{v})$ 在 V 的基 $\vec{v_1}, \cdots, \vec{v_n}$ 下对应 f 成规范形
$$u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \cdots - u_r^2.$$
若 $r < n$，则 $Q_f(\vec{v_n}) = 0$，与 ii) 矛盾；$\Rightarrow r = n$
若 $p < r = n$，则 $Q_f(\vec{v_n}) = -1$，与 ii) 矛盾；$\Rightarrow p = r = n$. Q.E.D.

- Sylvester 正定性准则
- 该准则给出了实二次型正定性判定的一个重要等价条件.
- 矩阵 A 的前 k 行前 k 列的行列式称为 A 的 k 阶顺序主子式.
- 实二次型 f 正定，当且仅当对应矩阵 A 的所有顺序主子式大于零 (6-9)

证：必要性：设 $Q_f(\vec{v})$ 在基 $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n$ 下成矩阵 A.

若 $Q_f(\vec{v})$ 正定，则存在可逆矩阵 P 使 $A = P^T I P$

$\Rightarrow \det A = \det P^T \cdot \det P = (\det P)^2 > 0$.

即正定矩阵的行列式大于零.

设 $Q_f(\vec{v})$ 在 $\mathrm{span}(\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k)$ 上的限制对应矩阵 A_k.

显然 A_k 正定，$\det A_k > 0$

$\Rightarrow$ A 的 k 阶顺序主子式大于零 $k = 1, 2, \cdots, n$.

充分性：对维数作归纳法：

i) $\dim V = 1$ 时，$f = a_{11}x_1^2$，$a_{11} > 0$，f 正定.

ii) 设 $\dim V = n$ 时，($\forall k, \det A_k > 0 \Rightarrow f$ 正定)

iii) 当 $\dim V = n+1$ 时，设 $Q_f(\vec{v})$ 在基 $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_{n+1}$ 下成矩阵 A.

由假设，$Q_f(\vec{v})$ 在 $\mathrm{span}(\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n)$ 上的限制对应矩阵 A_n 正定.

设 $U = \mathrm{span}(\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n)$，则 U 上有一组基 $\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_n$

使 $f(\vec{u}_i, \vec{u}_j) = \delta_{ij}$ $(i, j = 1, 2, \cdots, n)$

扩充为 V 的基 $\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_n, \vec{u}_{n+1}$.

令 $\vec{w} = \vec{u}_{n+1} - \sum_{i=1}^{n} f(\vec{u}_i, \vec{u}_{n+1})\vec{u}_i$

$$
\begin{aligned}
\text{于是 } f(\vec{u}_j, \vec{w}) &= f(\vec{u}_j, \vec{u}_{n+1}) - \sum_{i=1}^{n} f(\vec{u}_j, \vec{u}_i) f(\vec{u}_i, \vec{u}_{n+1}) \\
&= f(\vec{u}_j, \vec{u}_{n+1}) - f(\vec{u}_{n+1}, \vec{u}_j) \\
&= 0 \qquad (j = 1, 2, \cdots, n)
\end{aligned}
$$

$\Rightarrow Q_f(\vec{v})$ 在基 $\vec{u}_1, \cdots, \vec{u}_n, \vec{w}$ 下的矩阵为 $\mathrm{diag}(I_n, f(\vec{w}, \vec{w}))$

$\det \mathrm{diag}(I_n, f(\vec{w}, \vec{w})) > 0 \Rightarrow f(\vec{w}, \vec{w}) = d > 0$.

$\Rightarrow Q_f(\vec{v})$ 在基 $\vec{u_1}, \cdots, \vec{u_n}, \frac{1}{\sqrt{a}}\vec{w}$ 下的矩阵为 I_{n+1}.

即 $Q_f(\vec{v})$ 正定.（f 正定）. Q.E.D

- 类似正定二次型，其它实二次型有等价条件：
- 正定二次型 $f = u_1^2 + \cdots + u_n^2$

  维数 n = 秩 r = 正惯性指数 p
- 半正定二次型 $f = u_1^2 + \cdots + u_r^2$

  秩 r = 正惯性指数 p
- 负定二次型 $f = -u_1^2 - \cdots - u_n^2$

  维数 n = 秩 r = 负惯性指数 $r-p$（正惯性指数 $p=0$）
- 半负定二次型 $f = -u_1^2 - \cdots - u_r^2$

  秩 r = 负惯性指数 $r-p$

内积空间

7.1 内积与范数

- $\mathbb{R}^n$ 中的点积与范数
- 把 $\mathbb{R}^2$ 与 $\mathbb{R}^3$ 中的向量视为箭头，箭头的长度称为向量的范数，记为 $\|\vec{x}\|$

  $\forall \vec{x} = (x_1, x_2)^T \in \mathbb{R}^2$，$\|\vec{x}\| = \sqrt{x_1^2 + x_2^2}$ （勾股定理）

  $\forall \vec{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3$，$\|\vec{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2}$
- 以上定义易推广至 $\mathbb{R}^n$：

  $\|\vec{x}\| = \sqrt{\sum_{i=1}^n x_i^2}$

※ 向量范数 $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ 显然不是线性的，需要引入点积：

- 点积是 $\mathbb{R}^n$ 上的正定对称双线性形式：

  $\forall \vec{x}, \vec{y} \in \mathbb{R}^n$，$\vec{x} \cdot \vec{y} = \sum_{i=1}^n x_i y_i$
- 点积诱导范数：$\|\vec{x}\| = \sqrt{\vec{x} \cdot \vec{x}}$
- Euclid 空间
- 将 $\mathbb{R}^n$ 上的点积推广到实向量空间 V 上，定义内积是如下映射

  $<,> : V \times V \to \mathbb{R}$，$(\vec{v}, \vec{w}) \mapsto <\vec{v}, \vec{w}>$

  满足性质：

  i) 第一位置的线性：

  $\forall \lambda_1, \lambda_2 \in \mathbb{R}, \vec{v}_1, \vec{v}_2, \vec{w} \in V$，$<\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2, \vec{w}> = \lambda_1 <\vec{v}_1, \vec{w}> + \lambda_2 <\vec{v}_2, \vec{w}>$

  ii) 对称性：

  $\forall \vec{v}, \vec{w} \in V$，$<\vec{v}, \vec{w}> = <\vec{w}, \vec{v}>$

  iii) 正定性：

  $\forall \vec{v} \in V$，$<\vec{v}, \vec{v}> \geq 0$；$<\vec{v}, \vec{v}> = 0 \iff \vec{v} = 0$.

※ i) + ii) 可导出第二位置的线性.

- 内积是实向量空间上的正定对称双线性形式，定义了内积的实向量空间称为 Euclid 空间.
- $\mathbb{C}^n$ 与酉空间

- 在$\mathbb{C}$上，复数的绝对值是实数

$\forall z = x+iy$，$x, y \in \mathbb{R}$，$|z| = \sqrt{x^2+y^2} = \sqrt{z\bar{z}}$

- 若希望定义$\mathbb{C}^n$上的范数使值域为非负实数，绝对值是必要的：

$\forall \vec{z} \in \mathbb{C}^n$，$\|\vec{z}\| = \sqrt{\sum_{i=1}^{n} |z_i|^2} = \sqrt{\sum_{i=1}^{n} z_i \bar{z}_i}$

- 类似地定义$\mathbb{C}^n$上的点积：

$\forall \vec{w}, \vec{z} \in \mathbb{C}^n$，$\vec{w} \cdot \vec{z} = \sum_{i=1}^{n} w_i \bar{z}_i$ （或$\sum_{i=1}^{n} \bar{w}_i z_i$，易证相等）

- 将$\mathbb{C}^n$上的点积推广到复向量空间V上，定义内积是如下的映射：

$\langle , \rangle : V \times V \to \mathbb{C}$，满足性质：

i) 第一位置的线性：

$\forall \lambda_1, \lambda_2 \in \mathbb{C}, \vec{v}_1, \vec{v}_2, \vec{w} \in V$，$\langle \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2, \vec{w} \rangle = \lambda_1 \langle \vec{v}_1, \vec{w} \rangle + \lambda_2 \langle \vec{v}_2, \vec{w} \rangle$

ii) 共轭对称性：

$\forall \vec{v}, \vec{w} \in \mathbb{C}$，$\langle \vec{v}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{v} \rangle}$

iii) 正定性：

$\forall \vec{v} \in \mathbb{C}$，$\langle \vec{v}, \vec{v} \rangle \in \mathbb{R}$ 且 $\langle \vec{v}, \vec{v} \rangle \geq 0$；$\langle \vec{v}, \vec{v} \rangle = 0 \iff \vec{v} = 0$

※ 实数可视为虚部为零的复数，因此以上内积的定义也兼容实向量空间，可以作为内积的一般定义。

- 定义了内积的复向量空间称为酉空间。

▸ 内积与范数的性质

- $\forall \vec{v} \in V$，$\langle \vec{v}, 0 \rangle = \langle 0, \vec{v} \rangle = 0$ (7-1)

证：$\langle 0, \vec{v} \rangle = \langle 0, \vec{v} \rangle + \langle 0, \vec{v} \rangle \Rightarrow \langle 0, \vec{v} \rangle = 0$

$\langle \vec{v}, 0 \rangle = \overline{\langle 0, \vec{v} \rangle} = \bar{0} = 0$ Q.E.D

- （第二位置的共轭线性）

$\forall \lambda_1, \lambda_2 \in \mathbb{F}, \vec{u}, \vec{v}_1, \vec{v}_2 \in V$，$\langle \vec{u}, \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 \rangle = \bar{\lambda}_1 \langle \vec{u}, \vec{v}_1 \rangle + \bar{\lambda}_2 \langle \vec{u}, \vec{v}_2 \rangle$ (7-2)

证：$\langle \vec{u}, \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 \rangle = \overline{\langle \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2, \vec{u} \rangle} = \overline{\lambda_1 \langle \vec{v}_1, \vec{u} \rangle} + \overline{\lambda_2 \langle \vec{v}_2, \vec{u} \rangle}$

$= \bar{\lambda}_1 \langle \vec{u}, \vec{v}_1 \rangle = \bar{\lambda}_2 \langle \vec{u}, \vec{v}_2 \rangle$. Q.E.D

※ 该性质说明一般的内积不是双线性形式（算 Hermite 双线性形式）
这是为了保证 $\langle\vec{v},\vec{v}\rangle$ 是非负实数的特性.

- $\forall\vec{w}\in V$, $\langle\vec{u},\vec{w}\rangle=\langle\vec{v},\vec{w}\rangle$ 当且仅当 $\vec{u}=\vec{v}$. (7-3)

证：充分性：$\vec{u}=\vec{v}\Rightarrow\vec{u}-\vec{v}=0\Rightarrow 0=\langle 0,\vec{w}\rangle=\langle\vec{u}-\vec{v},\vec{w}\rangle$
$\Rightarrow\langle\vec{u},\vec{w}\rangle=\langle\vec{v},\vec{w}\rangle$

必要性：$\forall\vec{w}\in V$, $\langle\vec{u},\vec{w}\rangle=\langle\vec{v},\vec{w}\rangle\Rightarrow\langle\vec{u}-\vec{v},\vec{w}\rangle=0$
令 $\vec{w}=\vec{u}-\vec{v}$. $\Rightarrow\langle\vec{u}-\vec{v},\vec{u}-\vec{v}\rangle=0\Rightarrow\vec{u}-\vec{v}=0\Rightarrow\vec{u}=\vec{v}$ Q.E.D.

- 每种内积都诱导出一种范数：$\|\vec{v}\|=\sqrt{\langle\vec{v},\vec{v}\rangle}$
- $\|\vec{v}\|=0$ 当且仅当 $\vec{v}=0$
- $\forall\lambda\in\mathbb{F}$, $\|\lambda\vec{v}\|=|\lambda|\cdot\|\vec{v}\|$. (7-4)

证：$\|\lambda\vec{v}\|^2=\langle\lambda\vec{v},\lambda\vec{v}\rangle=\lambda\bar{\lambda}\langle\vec{v},\vec{v}\rangle=|\lambda|^2\|\vec{v}\|^2$
$\Rightarrow\|\lambda\vec{v}\|=|\lambda|\|\vec{v}\|$ Q.E.D.

▸ 内积与范数有关的定理

- Cauchy-Schwarz 不等式
$|\langle\vec{u},\vec{v}\rangle|\leqslant\|\vec{u}\|\cdot\|\vec{v}\|$ 等号成立当且仅当 $\vec{u}$ 与 $\vec{v}$ 线性相关. (7-5)

证：先在 Euclid 空间中证：
$0\leqslant\|\vec{u}-\lambda\vec{v}\|=\langle\vec{u}-\lambda\vec{v},\vec{u}-\lambda\vec{v}\rangle=\|\vec{u}\|^2-2\lambda\langle\vec{u},\vec{v}\rangle+\lambda^2\|\vec{v}\|^2$.
右式视为关于 λ 的二次函数，有 $\Delta=4\langle\vec{u},\vec{v}\rangle^2-4\|\vec{v}\|^2\|\vec{u}\|^2\leqslant 0$
$\Rightarrow|\langle\vec{u},\vec{v}\rangle|\leqslant\|\vec{u}\|\cdot\|\vec{v}\|$
注意函数极值在 $\lambda=\frac{\langle\vec{u},\vec{v}\rangle}{\|\vec{v}\|^2}$ 时取到.
在酉空间中，$0\leqslant\langle\vec{u}-\lambda\vec{v},\vec{u}-\lambda\vec{v}\rangle$
$$=\|\vec{u}\|^2-\lambda\langle\vec{v},\vec{u}\rangle-\bar{\lambda}\langle\vec{u},\vec{v}\rangle+|\lambda|^2\|\vec{v}\|^2.$$
代入 $\lambda=\frac{\langle\vec{u},\vec{v}\rangle}{\|\vec{v}\|^2}$ 得：
$$0\leqslant\|\vec{u}\|^2-\frac{|\langle\vec{u},\vec{v}\rangle|^2}{\|\vec{v}\|^2}-\frac{|\langle\vec{u},\vec{v}\rangle|^2}{\|\vec{v}\|^2}+\frac{|\langle\vec{u},\vec{v}\rangle|^2}{\|\vec{v}\|^4}\|\vec{v}\|^2$$
$$=\|\vec{u}\|^2-\frac{|\langle\vec{u},\vec{v}\rangle|^2}{\|\vec{v}\|^2}$$

$\Rightarrow |\langle\vec{u},\vec{v}\rangle| \leq \|\vec{u}\|\cdot\|\vec{v}\|$.

等号成立即 $\|\vec{u}-\lambda\vec{v}\|^2=0 \Leftrightarrow \vec{u}-\lambda\vec{v}=0 \Leftrightarrow \vec{u}$ 与 $\vec{v}$ 线性相关.

※ Cauchy-Schwarz不等式的应用：

i) $\forall x_1,\cdots,x_n,y_1,\cdots,y_n \in R$，$(\sum_{i=1}^n x_iy_i)^2 \leq \sum_{i=1}^n x_i^2\cdot\sum_{i=1}^n y_i^2$

ii) $\forall [a,b]$ 在连续函数 $f(x),g(x)$，

$$\left(\int_a^b f(x)g(x)dx\right)^2 \leq \left(\int_a^b f^2(x)dx\right)\left(\int_a^b g^2(x)dx\right)$$

iii) $|\langle\vec{u},\vec{v}\rangle| \leq \|\vec{u}\|\cdot\|\vec{v}\| \Rightarrow -1 \leq \frac{|\langle\vec{u},\vec{v}\rangle|}{\|\vec{u}\|\cdot\|\vec{v}\|} \leq 1$.

定义向量 $\vec{u}$ 与 $\vec{v}$ 的夹角 $\theta = \arccos\frac{|\langle\vec{u},\vec{v}\rangle|}{\|\vec{u}\|\cdot\|\vec{v}\|}$

- 正交性
- 向量 $\vec{u}$ 与 $\vec{v}$ 正交，若 $\langle\vec{u},\vec{v}\rangle=0$.
- 0 与任意向量正交.
- 两个正交的向量夹角 $\theta=\frac{\pi}{2}$（几何上可以说向量"垂直"）
- 勾股定理（两直角边平方和等于斜边平方）

若 $\vec{u}$ 与 $\vec{v}$ 正交，则 $\|\vec{u}+\vec{v}\|^2=\|\vec{u}\|^2+\|\vec{v}\|^2$ (7-6)

证：$\|\vec{u}+\vec{v}\|^2 = \langle\vec{u}+\vec{v},\vec{u}+\vec{v}\rangle = \langle\vec{u},\vec{u}\rangle+\langle\vec{u},\vec{v}\rangle+\langle\vec{v},\vec{u}\rangle+\langle\vec{v},\vec{v}\rangle$

$= \langle\vec{u},\vec{u}\rangle+\langle\vec{v},\vec{v}\rangle = \|\vec{u}\|^2+\|\vec{v}\|^2$ Q.E.D

※ 若 $\|\vec{u}+\vec{v}\|^2=\|\vec{u}\|^2+\|\vec{v}\|^2$，则有 $2Re\langle\vec{u},\vec{v}\rangle=0$.

即勾股定理的逆命题仅在Euclid空间成立. (7-7)

- 正交分解

把 $\vec{u}$ 向 $\vec{v}$ 的方向分解，即找到 λ，$\vec{w}$ 使得 $\vec{u}=\lambda\vec{v}+\vec{w}$ 且 $\langle\vec{v},\vec{w}\rangle=0$

$\Rightarrow \langle\vec{u}-\lambda\vec{v},\vec{v}\rangle=0$

$\Rightarrow \langle\vec{u},\vec{v}\rangle-\lambda\|\vec{v}\|^2=0 \Rightarrow \lambda=\frac{\langle\vec{u},\vec{v}\rangle}{\|\vec{v}\|^2}$

$\vec{w}=\vec{u}-\lambda\vec{v}=\vec{u}-\frac{\langle\vec{u},\vec{v}\rangle}{\|\vec{v}\|^2}\vec{v}$ (7-8)

※ Cauchy-Schwarz也可通过勾股定理+正交分解证明.

- 三角不等式（两边之和不小于第三边）

$$\|\vec{u}+\vec{v}\| \leqslant \|\vec{u}\| + \|\vec{v}\| \qquad (7\text{-}9)$$

证: $\|\vec{u}+\vec{v}\|^2 = \langle \vec{u}+\vec{v}, \vec{u}+\vec{v} \rangle = \langle \vec{u},\vec{u}\rangle + \langle \vec{v},\vec{v}\rangle + \langle \vec{u},\vec{v}\rangle + \overline{\langle \vec{u},\vec{v}\rangle}$

$$= \|\vec{u}\|^2 + \|\vec{v}\|^2 + 2Re\langle \vec{u},\vec{v}\rangle$$

$$\leqslant \|\vec{u}\|^2 + \|\vec{v}\|^2 + 2|\langle \vec{u},\vec{v}\rangle|$$

$$\leqslant \|\vec{u}\|^2 + \|\vec{v}\|^2 + 2\|\vec{u}\|\cdot\|\vec{v}\| \quad \text{(Cauchy-Schwarz不等式)}$$

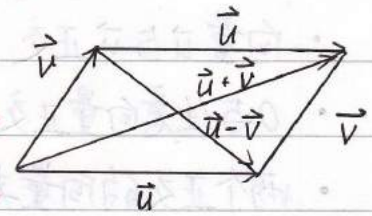
$$= (\|\vec{u}\| + \|\vec{v}\|)^2$$

$\Rightarrow \|\vec{u}+\vec{v}\| \leqslant \|\vec{u}\| + \|\vec{v}\|$

等号成立条件是 $\vec{u}$ 与 $\vec{v}$ "同向"。 Q.E.D.

- 平行四边形恒等式(对角线平方和等于四边平方和) (7-10)

$$\|\vec{u}+\vec{v}\|^2 + \|\vec{u}-\vec{v}\|^2 = 2(\|\vec{u}\|^2 + \|\vec{v}\|^2)$$



证: $\|\vec{u}+\vec{v}\|^2 + \|\vec{u}-\vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 + \langle \vec{u},\vec{v}\rangle + \langle \vec{v},\vec{u}\rangle$

$$+ \|\vec{u}\|^2 + \|\vec{v}\|^2 - \langle \vec{u},\vec{v}\rangle - \langle \vec{v},\vec{u}\rangle$$

$$= 2(\|\vec{u}\|^2 + \|\vec{v}\|^2)$$

Q.E.D.

▶ 赋范空间 & 度量空间

- 内积诱导范数: $\|\vec{v}\| = \sqrt{\langle \vec{v},\vec{v}\rangle}$, 而范数诱导度量: $d(\vec{u},\vec{v}) = \|\vec{u}-\vec{v}\|$
- 范数的定义可以不依赖内积, 满足如下性质的泛函 $\|\cdot\|: V \to R$ 都可以作为范数: (7-11)

  i) 正定性: $\forall \vec{v}\in V$, $\|\vec{v}\| \geqslant 0$; $\|\vec{v}\| = 0 \Leftrightarrow \vec{v} = 0$

  ii) 齐性: $\forall \lambda \in \mathbb{F}$, $\vec{v}\in V$, $\|\lambda\vec{v}\| = |\lambda|\cdot\|\vec{v}\|$

  iii) 三角不等式: $\forall \vec{v},\vec{w}\in V$, $\|\vec{v}+\vec{w}\| \leqslant \|\vec{v}\| + \|\vec{w}\|$

  定义了范数的向量空间称为赋范空间.
- 度量则是更加基本的结构(甚至不依赖向量空间), 度量函数 $d: S\times S \to R$ 满足如下性质: (7-12)

  i) 正定性: $\forall x,y\in S$, $d(x,y) > 0$; $d(x,x) = 0$

  ii) 对称性: $\forall x,y\in S$, $d(x,y) = d(y,x)$

(iii) 三角不等式：$\forall x,y,z\in S$，$d(x,y)\leq d(x,z)+d(z,y)$

- 若 $\|\cdot\|$ 是 V 上满足平行四边形恒等式的范数，则存在 V 上的一个内积 $\langle\ ,\ \rangle$ 使得 $\forall \vec{u}\in V$，$\|\vec{u}\|=\sqrt{\langle\vec{u},\vec{u}\rangle}$. (7-13)

（范数 + 平行四边形恒等式 $\xrightarrow{\text{诱导}}$ 内积）

证：先证实向量空间的情况.

定义 $f: V\times V\to\mathbb{R}$，$\forall \vec{u},\vec{v}\in V$，$f(\vec{u},\vec{v})=\frac{1}{4}(\|\vec{u}+\vec{v}\|^2-\|\vec{u}-\vec{v}\|^2)$

0. $f(\vec{u},\vec{u})=\frac{1}{4}(\|\vec{u}+\vec{u}\|^2+\|0\|^2)=\|\vec{u}\|^2$

$\Rightarrow \|\vec{u}\|=\sqrt{f(\vec{u},\vec{u})}$ 范数与内积关系满足

下面证 f 满足内积定义.

1. 由范数诱导的 f 自动满足正定性

2. 对称性：

$$f(\vec{u},\vec{v})=\frac{1}{4}(\|\vec{u}+\vec{v}\|^2-\|\vec{u}-\vec{v}\|^2)=\frac{1}{4}\left[\|\vec{v}+\vec{u}\|^2-(-1)^2\|\vec{v}-\vec{u}\|^2\right]$$
$$=f(\vec{v},\vec{u})$$

3. 第一位置的加性：

$\forall \vec{u}_1,\vec{u}_2,\vec{v}\in V$，$4[f(\vec{u}_1+\vec{u}_2,\vec{v})-f(\vec{u}_1,\vec{v})-f(\vec{u}_2,\vec{v})]$

$$=(\|u_1+u_2+v\|^2-\|u_1+u_2-v\|^2)+(\|u_1+v\|^2-\|u_1-v\|^2)+(\|u_2+v\|^2-\|u_2-v\|^2)$$

$$=\|u_1+u_2+v\|^2+(\|u_1-v\|^2+\|u_2-v\|^2)-\|u_1+u_2-v\|^2-(\|u_1+v\|^2+\|u_2+v\|^2)$$

$$=\|u_1+u_2+v\|^2+\frac{1}{2}(\|u_1+u_2-2v\|^2+\underline{\|u_1-u_2\|^2})-\|u_1+u_2-v\|^2-\frac{1}{2}(\|u_1+u_2+2v\|^2+\cancel{\|u_1-u_2\|^2})$$

$$=(\|u_1+u_2+v\|^2+\|v\|^2)+\frac{1}{2}\|u_1+u_2-2v\|^2-(\|u_1+u_2-v\|^2+\|v\|^2)-\frac{1}{2}\|u_1+u_2+2v\|^2$$

$$=\frac{1}{2}\|u_1+u_2+2v\|^2+\frac{1}{2}\|u_1+u_2\|^2+\frac{1}{2}\|u_1+u_2-2v\|^2$$
$$-\frac{1}{2}\|u_1+u_2\|^2-\frac{1}{2}\|u_1+u_2-2v\|^2-\frac{1}{2}\|u_1+u_2+2v\|^2$$

$$=0$$

$\Rightarrow f(\vec{u}_1+\vec{u}_2,\vec{v})=f(\vec{u}_1,\vec{v})+f(\vec{u}_2,\vec{v})$

4. 第一位置的齐性：（思路 $\mathbb{Z}^+\to\mathbb{Q}^+\to\mathbb{Q}\to\mathbb{R}$）

i) $\forall n\in\mathbb{Z}^+,\ \vec{u}\in V,\ f(n\vec{u},\vec{v})=f(\sum_{i=1}^{n}\vec{u},\vec{v})=\sum_{i=1}^{n}f(\vec{u},\vec{v})=nf(\vec{u},\vec{v})$

ii) $\forall m\in\mathbb{Z}^+,\ \vec{u}\in V,\ f(\vec{u},\vec{v})=f(m\cdot\frac{\vec{u}}{m},\vec{v})=mf(\frac{\vec{u}}{m},\vec{v})$

$$\Rightarrow f(\frac{\vec{u}}{m},\vec{v})=\frac{1}{m}f(\vec{u},\vec{v})$$

iii) $\forall r\in\mathbb{Q}^+,\ \exists m,n\in\mathbb{Z}^+$ s.t. $r=\frac{n}{m}$.

$\forall\vec{u}\in V,\ f(r\vec{u},\vec{v})=f(\frac{n}{m}\vec{u},\vec{v})=\frac{n}{m}f(\vec{u},\vec{v})=rf(\vec{u},\vec{v})$.

iv) $\forall\vec{u},\vec{v}\in V,\ f(-\vec{u},\vec{v})=\frac{1}{4}(\|-\vec{u}+\vec{v}\|^2-\|-\vec{u}-\vec{v}\|^2)$

$$=-\frac{1}{4}(\|\vec{u}+\vec{v}\|^2-\|\vec{u}-\vec{v}\|^2)=-f(\vec{u},\vec{v})$$

$\Rightarrow\forall r\in\mathbb{Q},\ f(r\vec{u},\vec{v})=rf(\vec{u},\vec{v})$.

接下来用有理数列逼近实数：

v) $\forall\lambda\in\mathbb{R}$，$\exists$ 有理数序列 $\{r_n\}$ s.t. $\lim_{n\to\infty}r_n=\lambda$.

$\forall\vec{u},\vec{v}\in V,\ \lambda f(\vec{u},\vec{v})=\lim_{n\to\infty}r_nf(\vec{u},\vec{v})=\lim_{n\to\infty}f(r_n\vec{u},\vec{v})$

$=\frac{1}{4}\lim_{n\to\infty}(\|r_n\vec{u}+\vec{v}\|^2-\|r_n\vec{u}-\vec{v}\|^2)$

下面只要证 $\lim_{n\to\infty}\|r_n\vec{u}+\vec{v}\|=\|\lambda\vec{u}+\vec{v}\|$：

由三角不等式易得 $\big|\|r_n\vec{u}+\vec{v}\|-\|\lambda\vec{u}+\vec{v}\|\big|\le\|r_n\vec{u}-\lambda\vec{u}\|=|r_n-\lambda|\cdot\|\vec{u}\|$

$\Rightarrow\lim_{n\to\infty}(\|r_n\vec{u}+\vec{v}\|-\|\lambda\vec{u}+\vec{v}\|)\le\lim_{n\to\infty}|r_n-\lambda|\cdot\|\vec{u}\|=0$

$\Rightarrow\lim_{n\to\infty}\|r_n\vec{u}+\vec{v}\|=\|\lambda\vec{u}+\vec{v}\|$，同理 $\lim_{n\to\infty}\|r_n\vec{u}-\vec{v}\|=\|\lambda\vec{u}-\vec{v}\|$

$\Rightarrow\lambda f(\vec{u},\vec{v})=\frac{1}{4}(\|\lambda\vec{u}+\vec{v}\|^2-\|\lambda\vec{u}-\vec{v}\|^2)=f(\lambda\vec{u},\vec{v})$.

这样就证明了实向量空间的情况.

下面考虑复向量空间.

将上一部分的 f 记为 f_R. $\forall\vec{u},\vec{v}\in V,\ f_R(\vec{u},\vec{v})=\frac{1}{4}(\|\vec{u}+\vec{v}\|^2-\|\vec{u}-\vec{v}\|^2)$

定义 $f(\vec{u},\vec{v})=f_R(\vec{u},\vec{v})+i\cdot f_R(\vec{u},i\vec{v})$　（怎么想到的?!）

0. $f(\vec{u},\vec{u})=f_R(\vec{u},\vec{u})+if_R(\vec{u},i\vec{u})$

$$=\|\vec{u}\|^2+\frac{i}{4}(\|\vec{u}+i\vec{u}\|^2-\|\vec{u}-i\vec{u}\|^2)$$

$$=\|\vec{u}\|^2+\frac{i}{4}(|1+i|^2-|1-i|^2)\|\vec{u}\|^2$$

$$=\|\vec{u}\|^2$$

$$\Rightarrow \|\vec{u}\| = \sqrt{f(\vec{u},\vec{u})}$$

下面证 $f(\vec{u},\vec{v})$ 满足内积定义：

1. 正定性. 自动满足

2. 共轭对称性.

$$
\begin{aligned}
f(\vec{u},\vec{v}) &= f_R(\vec{u},\vec{v}) + i f_R(\vec{u}, i\vec{v}) \\
&= f_R(\vec{v},\vec{u}) + i f_R(i\vec{v},\vec{u}) \\
&= \overline{f_R(\vec{v},\vec{u})} + \frac{i}{4}\left(\|i\vec{v}+\vec{u}\|^2 - \|i\vec{v}-\vec{u}\|^2\right) \\
&= \overline{f_R(\vec{v},\vec{u})} + \frac{i}{4}\left[\|i(v-i\vec{u})\|^2 - \|(-i)(\vec{v}+i\vec{u})\|^2\right] \\
&= \overline{f_R(\vec{v},\vec{u})} + \overline{\frac{i}{4}\left(\|\vec{v}+i\vec{u}\|^2 - \|\vec{v}-i\vec{u}\|^2\right)} \\
&= \overline{f_R(\vec{v},\vec{u})} + \overline{i f_R(\vec{v}, i\vec{u})} \\
&= \overline{f(\vec{v},\vec{u})}
\end{aligned}
$$

3. 第一位置的加性.

$$
\begin{aligned}
f(\vec{u}_1+\vec{u}_2,\vec{v}) &= f_R(\vec{u}_1+\vec{u}_2,\vec{v}) + i f_R(\vec{u}_1+\vec{u}_2, i\vec{v}) \\
&= f_R(\vec{u}_1,\vec{v}) + i f_R(\vec{u}_1, i\vec{v}) + f_R(\vec{u}_2,\vec{v}) + i f_R(\vec{u}_2+i\vec{v}) \\
&= f(\vec{u}_1,\vec{v}) + f(\vec{u}_2,\vec{v})
\end{aligned}
$$

4. 第一位置的齐性.

i) $\forall \lambda \in \mathbb{R}$,
$$
\begin{aligned}
f(\lambda\vec{u},\vec{v}) &= f_R(\lambda\vec{u},\vec{v}) + i f_R(\lambda\vec{u}, i\vec{v}) \\
&= \lambda f_R(\vec{u},\vec{v}) + i\lambda f_R(\vec{u}, i\vec{v}) \\
&= \lambda f(\vec{u},\vec{v})
\end{aligned}
$$

ii)
$$
\begin{aligned}
f(i\vec{u},\vec{v}) &= f_R(i\vec{u},\vec{v}) + i f_R(i\vec{u}, i\vec{v}) \\
&= \frac{1}{4}\left(\|i\vec{u}+\vec{v}\|^2 - \|i\vec{u}-\vec{v}\|^2 + i\|i\vec{u}+i\vec{v}\|^2 - i\|i\vec{u}-i\vec{v}\|^2\right) \\
&= \frac{1}{4}\left(i\|\vec{u}+\vec{v}\|^2 - i\|\vec{u}-\vec{v}\|^2 + i\|\vec{u}+i\vec{v}\|^2 - i\|\vec{u}-i\vec{v}\|^2\right) \\
&= i f(\vec{u},\vec{v})
\end{aligned}
$$

$\Rightarrow \forall z \in \mathbb{C}$, $\exists a, b \in \mathbb{R}$ s.t. $z = a+bi$

$$f(z\vec{u},\vec{v}) = f((a+bi)\vec{u},\vec{v}) = (a+bi) f(\vec{u},\vec{v}) = z f(\vec{u},\vec{v}).$$

这样就证明了复向量空间的情况. Q.E.D.

※ 以下两个恒等式较为常用:

Euclid空间中

$$\langle\vec{u},\vec{v}\rangle=\frac{1}{4}\left(\|\vec{u}+\vec{v}\|-\|\vec{u}-\vec{v}\|\right) \tag{7-14}$$

酉空间中

$$\langle\vec{u},\vec{v}\rangle=\frac{1}{4}\left(\|\vec{u}+\vec{v}\|-\|\vec{u}-\vec{v}\|+i\|\vec{u}+i\vec{v}\|-i\|\vec{u}-i\vec{v}\|\right) \tag{7-15}$$

$$=\frac{1}{4}\sum_{n=1}^{4}i^n\|\vec{u}+i^n\vec{v}\|^2 \tag{7-16}$$

7.2 标准正交基

▸ 正交组 & 标准正交组

- V上向量组 $\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_n$ 是正交组, 若向量两两正交.

  即 $\forall j\neq k$, $\langle\vec{v}_j,\vec{v}_k\rangle=0$.

- 广义勾股定理

  若 $\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_n$ 是V上正交组, 则 $\forall\lambda_1,\lambda_2,\cdots,\lambda_n\in\mathbb{F}$, 有:

$$\left\|\sum_{j=1}^{n}\lambda_j\vec{v}_j\right\|^2=\sum_{j=1}^{n}|\lambda_j|^2\|\vec{v}_j\|^2 \tag{7-17}$$

证: $\left\|\sum_{j=1}^{n}\lambda_j\vec{v}_j\right\|^2=\left\langle\sum_{j=1}^{n}\lambda_j\vec{v}_j,\sum_{k=1}^{n}\lambda_k\vec{v}_k\right\rangle=\sum_{j=1}^{n}\sum_{k=1}^{n}\lambda_j\overline{\lambda_k}\langle\vec{v}_j,\vec{v}_k\rangle\delta_{j,k}$ ← 正交性

$=\sum_{j=1}^{n}\lambda_j\overline{\lambda_j}\langle\vec{v}_j,\vec{v}_j\rangle=\sum_{j=1}^{n}|\lambda_j|^2\|\vec{v}_j\|^2$ Q.E.D.

- 正交组是线性无关组. (7-18)

证: 设 $\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_n$ 是V上正交组.

设 $\lambda_1,\lambda_2,\cdots,\lambda_n\in\mathbb{F}$ s.t. $\sum_{j=1}^{n}\lambda_j\vec{v}_j=0$

$\Rightarrow\left\|\sum_{j=1}^{n}\lambda_j\vec{v}_j\right\|^2=\sum_{j=1}^{n}|\lambda_j|^2\|\vec{v}_j\|^2=0$

由范数非负得 $\lambda_1=\lambda_2=\cdots=\lambda_n=0$, 即 $\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_n$ 线性无关 Q.E.D.

- V上向量组 $\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n$ 是标准正交组, 若向量两两正交且范数等于1.

  即 $\forall j,k$ $\langle\vec{e}_j,\vec{e}_k\rangle=\delta_{j,k}$

▸ 正交基 & 标准正交基

- V中长度为$\dim V$的正交组是V的基，称为V的正交基. (7-19)
- Fourier分解（有限维简化版） (7-20)

  设$\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_n$是$V$的正交基

  $\forall \vec{v}\in V$，设$\vec{v}=\sum_{j=1}^{n}\lambda_j\vec{v}_j$.

  $\Rightarrow \langle\vec{v},\vec{v}_k\rangle=\langle\sum_{j=1}^{n}\lambda_j\vec{v}_j,\vec{v}_k\rangle=\sum_{j=1}^{n}\lambda_j\langle\vec{v}_j,\vec{v}_k\rangle\delta_{j,k}=\lambda_k\langle\vec{v}_k,\vec{v}_k\rangle=\lambda_k\|\vec{v}_k\|^2$

  $\Rightarrow \lambda_k=\dfrac{\langle\vec{v},\vec{v}_k\rangle}{\|\vec{v}_k\|^2}$

  即 $\vec{v}=\sum_{j=1}^{n}\dfrac{\langle\vec{v},\vec{v}_j\rangle}{\|\vec{v}_j\|^2}\vec{v}_j$.

  特别地，若$\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n$是$V$的标准正交基

  $\|\vec{e}_j\|=1$. $\forall\vec{v}\in V$，$\vec{v}=\sum_{j=1}^{n}\langle\vec{v},\vec{e}_j\rangle\vec{e}_j$.

  且有 $\|\vec{v}\|^2=\sum_{j=1}^{n}|\langle\vec{v},\vec{e}_j\rangle|^2$

- Parseval等式 (7-21)

  根据向量在标准正交基下的分解，内积的计算可以相应简化：

  设$\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n$是$V$的标准正交基，

  $\forall\vec{u},\vec{v}\in V$，$\langle\vec{u},\vec{v}\rangle=\left\langle\sum_{j=1}^{n}\langle\vec{u},\vec{e}_j\rangle\vec{e}_j,\sum_{k=1}^{n}\langle\vec{v},\vec{e}_k\rangle\vec{e}_k\right\rangle$

  $=\sum_{j=1}^{n}\sum_{k=1}^{n}\langle\vec{u},\vec{e}_j\rangle\overline{\langle\vec{v},\vec{e}_k\rangle}\langle\vec{e}_j,\vec{e}_k\rangle=\sum_{j=1}^{n}\sum_{k=1}^{n}\delta_{j,k}\langle\vec{u},\vec{e}_j\rangle\overline{\langle\vec{v},\vec{e}_k\rangle}$

  $\Rightarrow \langle\vec{u},\vec{v}\rangle=\sum_{j=1}^{n}\langle\vec{u},\vec{e}_j\rangle\overline{\langle\vec{v},\vec{e}_j\rangle}$

  这个结果相当于$\mathbb{C}^n$中的内积：$\vec{x}\cdot\vec{y}=\sum_{j=1}^{n}x_j\overline{y_j}$

- Gram-Schmidt正交化过程. (7-22)

  该算法可以将一个线性无关组化为一个标准正交组，且仍张成原空间.

  设$\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_m$是$V$中的线性无关组.

  令 $\vec{e}_1=\dfrac{\vec{v}_1}{\|\vec{v}_1\|}$；$\vec{e}_2=\dfrac{\vec{v}_2-\langle\vec{v}_2,\vec{e}_1\rangle\vec{e}_1}{\|\vec{v}_2-\langle\vec{v}_2,\vec{e}_1\rangle\vec{e}_1\|}$；$\cdots$；

  $\vec{e}_j=(\vec{v}_j-\sum_{k=1}^{j-1}\langle\vec{v}_j,\vec{e}_k\rangle\vec{e}_k)/\|\vec{v}_j-\sum_{k=1}^{j-1}\langle\vec{v}_j,\vec{e}_k\rangle\vec{e}_k\|$，$j=2,3,\cdots,m$.

  则$\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_m$是标准正交组，且$\mathrm{span}(\vec{v}_1,\cdots,\vec{v}_m)=\mathrm{span}(\vec{e}_1,\cdots,\vec{e}_m)$.

证：先从正交投影的角度直观理解该过程：

$\vec{v}_j=\vec{w}_j+\sum_{k=1}^{j-1}\langle\vec{v}_j,\vec{e}_k\rangle\vec{e}_k$.（$\vec{v}_j$投影分解到$\mathrm{span}(\vec{e}_1,\cdots,\vec{e}_{j-1})$上）

KOKUYO

则有 $\forall l=1,2,\cdots,j-1$. $\langle\vec{w_j},\vec{e_l}\rangle=\langle\vec{v_j}-\sum_{k=1}^{j-1}\langle\vec{v_j},\vec{e_k}\rangle e_k,\vec{e_l}\rangle$

$=\langle\vec{v_j},\vec{e_l}\rangle-\sum_{k=1}^{j-1}\langle\vec{v_j},\vec{e_k}\rangle\langle\vec{e_k},\vec{e_l}\rangle$

$=\langle\vec{v_j},\vec{e_l}\rangle-\langle\vec{v_j},\vec{e_l}\rangle=0$.

即 $\vec{w_j}$ 与 $\{\vec{e_1},\cdots,\vec{e_{j-1}}\}$ 整个组正交，则 $\vec{e_j}=\frac{\vec{w_j}}{\|\vec{w_j}\|}$ 与标准正交组 $\{\vec{e_1},\cdots,\vec{e_{j-1}}\}$ 的并 $\{\vec{e_1},\cdots,\vec{e_j}\}$ 仍为标准正交组.

下面给出严谨证明：

对 j 作归纳法.

i) $j=1$ 时，显然 $\mathrm{span}(\vec{v_1})=\mathrm{span}(\vec{e_1})$

ii) 假设 $1<j<m$ 时已有 $\mathrm{span}(\vec{v_1},\cdots,\vec{v_{j-1}})=\mathrm{span}(\vec{e_1},\cdots,\vec{e_{j-1}})$.

iii) $\vec{e_j}=(\vec{v_j}-\sum_{k=1}^{j-1}\langle\vec{v_j},\vec{e_k}\rangle\vec{e_k})/\|\vec{v_j}-\sum_{k=1}^{j-1}\langle\vec{v_j},\vec{e_k}\rangle\vec{e_k}\|$

$\Rightarrow\|\vec{e_j}\|=1$ 且前述已证明 $\forall l=1,\cdots,j-1$, $\langle\vec{e_j},\vec{e_l}\rangle=1$

$\Rightarrow\vec{e_1},\vec{e_2},\cdots,\vec{e_j}$ 是标准正交组

$\vec{v_j}\in\mathrm{span}(\vec{e_1},\cdots,\vec{e_j})\Rightarrow\mathrm{span}(\vec{v_1},\cdots,\vec{v_j})\subset\mathrm{span}(\vec{e_1},\cdots,\vec{e_j})$

又 $\dim\mathrm{span}(\vec{v_1},\cdots,\vec{v_j})=\dim\mathrm{span}(\vec{e_1},\cdots,\vec{e_j})=j$

$\Rightarrow\mathrm{span}(\vec{v_1},\cdots,\vec{v_j})=\mathrm{span}(\vec{e_1},\cdots,\vec{e_j})$. Q.E.D.

例：求 $R[x]_2$（或 $\mathcal{P}_2(R)$）的一组标准正交基

定义内积 $\langle p,q\rangle=\int_{-1}^{1}p(x)q(x)dx$

解：对基 $1,x,x^2$ 应用Gram-Schmidt过程：

e_1: $\|1\|^2=\int_{-1}^{1}1^2dx=2\Rightarrow\|1\|=\sqrt{2}\Rightarrow e_1=\frac{\sqrt{2}}{2}$

e_2: $x-\langle x,e_1\rangle e_1=x-\frac{\sqrt{2}}{2}\int_{-1}^{1}\frac{\sqrt{2}}{2}x\,dx=x$

$\|x\|^2=\int_{-1}^{1}x^2dx=\frac{2}{3}\Rightarrow\|x\|=\frac{\sqrt{6}}{3}$

$\Rightarrow e_2=\frac{x}{\|x\|}=\frac{\sqrt{6}}{2}x$

e_3: $x^2-\langle x^2,e_1\rangle e_1-\langle x^2,e_2\rangle e_2$

$=x^2-\frac{\sqrt{2}}{2}\int_{-1}^{1}\frac{\sqrt{2}}{2}x^2dx-\frac{\sqrt{6}}{2}x\int_{-1}^{1}x^2\cdot\frac{\sqrt{6}}{2}x\,dx$

$=x^2-\frac{1}{3}$

$\|x^2-\frac{1}{3}\| = \sqrt{\int_{-1}^{1}(x^2-\frac{1}{3})^2dx} = \frac{8}{45}$

$\Rightarrow e_3 = x^2-\frac{1}{3}/\|x^2-\frac{1}{3}\| = \frac{3\sqrt{10}}{4}(x^2-\frac{1}{3})$

即 $\mathbb{R}[x]_2$ 的一组标准正交基为 $\frac{\sqrt{2}}{2}$，$\frac{\sqrt{6}}{2}x$，$\frac{3\sqrt{10}}{4}(x^2-\frac{1}{3})$

- 有限维内积空间必有标准正交基 (7-23)
- 标准正交组可扩充为标准正交基. (7-24)

以上两个表述是 Gram-Schmidt 过程的直接推论.

- Schur 定理

对有限维酉空间 V 内算子 T，存在 V 的标准正交基使得 T 成上三角矩阵 (7-25)

证：由5-22，$\exists$ V 的基 $\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_n$ 使 T 成上三角矩阵.

对 $\vec{v}_1,\cdots,\vec{v}_n$ 应用 Gram-Schmidt 过程得 V 的标准正交基 $\vec{e}_1,\cdots,\vec{e}_n$.

且有 $\forall j=1,2,\cdots,n$，$span(\vec{v}_1,\cdots,\vec{v}_j) = span(\vec{e}_1,\cdots,\vec{e}_j)$

由于 $\forall j$，T 在 $span(\vec{v}_1,\cdots,\vec{v}_j)$ 内不变

$\Rightarrow$ T 在 $span(\vec{e}_1,\cdots,\vec{e}_j)$ 内不变 $\Rightarrow$ T 在 $\vec{e}_1,\cdots,\vec{e}_n$ 下成上三角形. Q.E.D

- Riesz 表示定理 (有限维简化版) (7-26

$\forall$ 有限维内积空间 V 上的线性泛函 φ，$\exists$ 唯一的 $\vec{u}\in V$ s.t. $\forall \vec{v}\in V$，$\varphi(\vec{v}) = \langle\vec{v},\vec{u}\rangle$.

(用8-1对偶空间的语言说，内积是从 V 到 V' 的一个自然同构)

证：存在性：设 $\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n$ 是 V 的标准正交基

$$\forall \vec{v}\in V,\ \varphi(\vec{v}) = \varphi\left(\sum_{j=1}^{n}\langle\vec{v},\vec{e}_j\rangle\vec{e}_j\right) = \sum_{j=1}^{n}\langle\vec{v},\vec{e}_j\rangle\varphi(\vec{e}_j) = \left\langle\vec{v},\sum_{j=1}^{n}\overline{\varphi(\vec{e}_j)}\,\vec{e}_j\right\rangle$$

取 $\vec{u} = \sum_{j=1}^{n}\overline{\varphi(\vec{e}_j)}\,\vec{e}_j$ 即满足条件

唯一性：设 $\vec{u}_1,\vec{u}_2\in V$ s.t. $\forall\vec{v}\in V$，$\varphi(\vec{v}) = \langle\vec{v},\vec{u}_1\rangle = \langle\vec{v},\vec{u}_2\rangle$

$\Rightarrow 0 = \langle\vec{v},\vec{u}_1\rangle - \langle\vec{v},\vec{u}_2\rangle = \langle\vec{v},\vec{u}_1-\vec{u}_2\rangle$，$\forall\vec{v}\in V$

取 $\vec{v} = \vec{u}_1-\vec{u}_2$ 得 $\vec{u}_1-\vec{u}_2 = 0 \Rightarrow \vec{u}_1 = \vec{u}_2$. Q.E.D

例：求 $u(t) \in \mathbb{R}[x]_2$（或 $\mathcal{P}_2(\mathbb{R})$）使得 $\forall p(t) \in \mathbb{R}[x]_2$，均有

$$\int_{-1}^{1} p(t)\cos\pi t\,dt = \int_{-1}^{1} p(t)u(t)\,dt$$

解：令线性泛函 $\varphi(p) = \int_{-1}^{1} p(t)\cos\pi t\,dt$

应用以上证明中的公式 $u = \sum_{j=1}^{n} \overline{\varphi(e_j)}\,e_j$

取标准正交基 $e_1 = \frac{\sqrt{2}}{2}$，$e_2 = \frac{\sqrt{6}}{2}x$，$e_3 = \frac{3\sqrt{10}}{4}(x^2-\frac{1}{3})$

$$\text{得 } u(x) = \frac{\sqrt{2}}{2}\int_{-1}^{1}\frac{\sqrt{2}}{2}\cos\pi t\,dt + \frac{\sqrt{6}}{2}x\int_{-1}^{1}\frac{\sqrt{6}}{2}t\cos\pi t\,dt$$
$$+ \frac{3\sqrt{10}}{4}(x^2-\frac{1}{3})\int_{-1}^{1}\frac{3\sqrt{10}}{4}(t^2-\frac{1}{3})\cos\pi t\,dt$$
$$= -\frac{45}{2\pi^2}(x^2-\frac{1}{3})$$

※ Riesz表示定理对无限维内积空间未必成立

7.3 正交补与正交投影

▶ 正交补

- 向量 $\vec{v}$ 与子集 U 正交：$\vec{v} \perp U \iff \forall \vec{u} \in U$，$\langle \vec{v}, \vec{u} \rangle = 0$.
- 子集 U 的正交补 $U^\perp$ 是所有与 U 正交的向量的集合.

  $U^\perp = \{\vec{v} \in V \mid \vec{v} \perp U\}$.

※ $\mathbb{R}^3$ 中，若 U 是过原点直线，则 $U^\perp$ 是垂直于 U 且过原点的平面.

- $U^\perp$ 是 V 的子空间 (7-27)

证：i) $\forall \vec{u} \in U$，$\langle 0, \vec{u} \rangle = 0 \Rightarrow 0 \perp U \Rightarrow 0 \in U^\perp$

ii) 加法封闭：$\forall \vec{v}, \vec{w} \in U^\perp$，有 $\forall \vec{u} \in U$，$\langle \vec{v}+\vec{w}, \vec{u} \rangle = \langle \vec{v}, \vec{u} \rangle + \langle \vec{w}, \vec{u} \rangle$

$= 0 + 0 = 0 \quad \Rightarrow \vec{v}+\vec{w} \in U^\perp$

iii) 数乘封闭：$\forall \vec{v} \in U^\perp$，$\lambda \in F$，有 $\forall \vec{u} \in U$，$\langle \lambda\vec{v}, \vec{u} \rangle = \lambda\langle \vec{v}, \vec{u} \rangle = 0$.

$\Rightarrow \lambda\vec{v} \in U^\perp$. Q.E.D.

- $U \cap U^\perp \subset \{0\}$ (7-28)

证：$\forall \vec{v} \in U \cap U^\perp$，$\langle \vec{v}, \vec{v} \rangle = 0 \Rightarrow \vec{v} = 0 \Rightarrow U \cap U^\perp \subset \{0\}$ Q.E.D.

- 若 U 是 V 的有限维子空间，则有 $V = U \oplus U^\perp$ (7-29)

证：设 U 有标准正交基 $\vec{e_1},\vec{e_2},\cdots,\vec{e_m}$.

$\forall \vec{v}\in V$. 令 $\vec{u}=\sum_{j=1}^{m}\langle\vec{v},\vec{e_j}\rangle\vec{e_j}$, $\vec{w}=\vec{v}-\vec{u_n}=\vec{v}-\sum_{j=1}^{m}\langle\vec{v},\vec{e_j}\rangle\vec{e_j}$

于是 $\forall k=1,2,\cdots,m$. $\langle\vec{w},\vec{e_k}\rangle=\langle\vec{v},\vec{e_k}\rangle-\sum_{j=1}\langle\vec{v},\vec{e_j}\rangle\langle\vec{e_j},\vec{e_k}\rangle=0$.

$\Rightarrow \vec{w}\perp span(\vec{e_1},\vec{e_2},\cdots,\vec{e_m})=U \Rightarrow \vec{w}\in U^\perp$

$\Rightarrow \forall\vec{v}\in V$, $\vec{v}=\vec{u}+\vec{w}$, 其中 $\vec{u}\in U$, $\vec{w}\in U^\perp \Rightarrow V=U+U^\perp$

又有 $U\cap U^\perp=\{0\}\Rightarrow V=U\oplus U^\perp$ Q.E.D.

※ 若 V 有限维且 U 是 V 的子空间，则 $\dim V=\dim U+\dim U^\perp$ (7-30)

- 若 U 是 V 的有限维子空间，则 $U=(U^\perp)^\perp$ (7-31)

证：$\forall\vec{u}\in U$, 有 $\forall\vec{v}\in U^\perp$, $\langle\vec{u},\vec{v}\rangle=0 \Rightarrow \vec{u}\perp U^\perp\Rightarrow\vec{u}\in(U^\perp)^\perp$

$\Rightarrow U\subset(U^\perp)^\perp$

$\forall\vec{v}\in(U^\perp)^\perp$, 有 $\vec{v}=\vec{u}+\vec{w}$, 其中 $\vec{u}\in U\subset(U^\perp)^\perp$, $\vec{w}\in U^\perp$

$\Rightarrow\vec{v}-\vec{u}\in(U^\perp)^\perp \Rightarrow \vec{w}=\vec{v}-\vec{u}\in U^\perp\cap(U^\perp)^\perp\Rightarrow\vec{w}=0$, $\vec{v}=\vec{u}$.

$\Rightarrow (U^\perp)^\perp\subset U$.

于是有 $U=(U^\perp)^\perp$ Q.E.D

※ 7-29与7-31在无限维子空间中不成立.

（分析没学好，反例不会举）

▸ 正交投影

- 设 U 是 V 的有限维子空间，定义 V 到 U 上的正交投影 $P_U\in\mathcal{L}(V,U)$

$\forall\vec{v}=\vec{u}+\vec{w}\in V$, 其中 $\vec{u}\in U$, $\vec{w}\in U^\perp$, 有 $P_U(\vec{v})=\vec{u}$

※ 该定义显然依赖于 $V=U\oplus U^\perp$.

- 正交投影的性质 (7-32)

i) $P_U\in\mathcal{L}(V)$

ii) $\forall\vec{u}\in U$. $P_U(\vec{u})=\vec{u}$

iii) $\forall\vec{w}\in U^\perp$, $P_U(\vec{w})=0$

iv) $Ran P_U=U$

v) $Null\, P_U = U^{\perp}$

vi) $V = Null\, P_U \oplus Ran\, P_U$

vii) $P_U^2 = P_U$

viii) $\|P_U(\vec{v})\| \leqslant \|\vec{v}\|$

ix) 若 $\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_m$ 是 U 的标准正交基

则 $\forall \vec{v} \in V$, $P_U(\vec{v}) = \sum_{j=1}^{m} <\vec{v}, \vec{e}_j> \vec{e}_j$

证: i) $\forall \vec{v}_1, \vec{v}_2 \in V$

设 $\vec{v}_1 = \vec{u}_1 + \vec{w}_1$, $\vec{v}_2 = \vec{u}_2 + \vec{w}_2$ 其中 $\vec{u}_1, \vec{u}_2 \in U$, $\vec{w}_1, \vec{w}_2 \in U^{\perp}$.

$P_U(\vec{v}_1 + \vec{v}_2) = \vec{u}_1 + \vec{u}_2 = P_U \vec{v}_1 + P_U \vec{v}_2$.

$\forall \vec{v} \in V$, $\lambda \in \mathbb{F}$. $\lambda \vec{v} = \lambda \vec{u} + \lambda \vec{w}$ 其中 $\lambda \vec{u} \in U$, $\lambda \vec{w} \in U^{\perp}$.

$P_U(\lambda \vec{v}) = \lambda \vec{u} = \lambda P_U(\vec{v})$.

即 P_U 是线性的. $P_U \in \mathcal{L}(V)$.

ii) 与 iii) 由定义显然

iv) 定义隐含 $Ran\, P_U \subset U$. 由 ii) $U \subset Ran\, P_U \Rightarrow Ran\, P_U = U$

v) 由 iii) $U^{\perp} \subset Null\, P_U$

$\forall \vec{v} \in Null\, P_U$, $\vec{v} = 0 + \vec{v}$ 其中 $0 \in U$, $\vec{v} \in U^{\perp} \Rightarrow Null\, P_U \subset U^{\perp}$

$\Rightarrow Null\, P_U = U^{\perp}$

vi) $V = U \oplus U^{\perp}$, $Null\, P_U = U^{\perp}$, $Ran\, P_U = U^{\perp}$

$\Rightarrow V = Null\, P_U \oplus Ran\, P_U$

vii) $\forall \vec{v} \in V$, $\vec{v} = \vec{u} + \vec{w}$.

$P_U^2(\vec{v}) = P_U(P_U(\vec{v})) = P_U(\vec{u}) = \vec{u} = P_U(\vec{v})$

$\Rightarrow P_U^2 = P_U$.

viii) $\|P_U(\vec{v})\|^2 = \|\vec{u}\|^2 \leqslant \|\vec{u}\|^2 + \|\vec{w}\|^2 = \|\vec{v}\|^2$

ix) $\forall \vec{v} \in V$, $\vec{v} = \vec{u} + \vec{w}$

其中 $\vec{u} = \sum_{j=1}^{m} <\vec{v}, \vec{e}_j> \vec{e}_j \in U$, $\vec{w} = \vec{v} - \sum_{j=1}^{m} <\vec{v}, \vec{e}_j> \vec{e}_j \in U^{\perp}$

$\Rightarrow P_U(\vec{v}) = \vec{u} = \sum_{j=1}^{m} \langle \vec{v}, \vec{e_j} \rangle \vec{e_j}$ Q.E.D.

- 极小化 & 最小二乘法

- 到子空间的最小距离

设U是V的有限维子空间，$\forall \vec{v} \in V, \vec{u} \in U$，

$\|\vec{v} - P_U(\vec{v})\| \leq \|\vec{v} - \vec{u}\|$

等号成立当且仅当 $P_U(\vec{v}) = \vec{u}$ (7-33)

证：$\vec{v} - P_U(\vec{v}) \in U^{\perp}$，$P_U(\vec{v}) - \vec{u} \in U$.

$\Rightarrow \langle \vec{v} - P_U(\vec{v}), P_U(\vec{v}) - \vec{u} \rangle = 0$

$$\begin{aligned}\Rightarrow \|\vec{v} - P_U(\vec{v})\|^2 &\leq \|\vec{v} - P_U(\vec{v})\|^2 + \|P_U(\vec{v}) - \vec{u}\|^2 \\ &= \|[\vec{v} - P_U(\vec{v})] + [P_U(\vec{v}) - \vec{u}]\|^2 \\ &= \|\vec{v} - \vec{u}\|^2\end{aligned}$$

其中 $\|\vec{v} - P_U(\vec{v})\| = \|\vec{v} - \vec{u}\| \Leftrightarrow \|P_U(\vec{v}) - \vec{u}\|^2 = 0 \Leftrightarrow \vec{u} = P_U(\vec{v})$. Q.E.D

- 最小二乘法.

第一章研究过矩阵方程 $A\vec{x} = \vec{b}$ 的解的结构.

$A\vec{x} = \vec{b}$ 有解当且仅当 $\vec{b} \in \mathrm{Ran}A$.

很多时候 $\vec{b} \notin \mathrm{Ran}A$，这时仍要得到一个近似解使误差最小，

即 $\|A\vec{x} - \vec{b}\|$ 取极小值（取Euclid内积，点积，对应Euclid距离）

由7-33结论，$A\vec{x} = P_{\mathrm{Ran}A}(\vec{b})$ 时 $\|A\vec{x} - \vec{b}\|$ 最小. (7-34)

- 一个直观的算法

i) 写出RanA的一组基（见4-25）

ii) 应用Gram-Schmidt过程得RanA的标准正交基 $\vec{e_1}, \vec{e_2}, \cdots, \vec{e_r}$

iii) $A\vec{x} = P_{\mathrm{Ran}A}(\vec{b}) = \sum_{j=1}^{r} \langle \vec{b}, \vec{e_j} \rangle \vec{e_j}$

iv) 若A满秩，用行化简法可得唯一的最小二乘解 $\vec{x}$.

- 正规方程

$\forall A \in \mathbb{R}^{m,n}$，$\vec{x} \in \mathbb{R}^n$，$\vec{b} \in \mathbb{R}^m$，若 $A\vec{x} = P_{\mathrm{Ran}A}(\vec{b})$，则有 $A^TA\vec{x} = A^T\vec{b}$ (7-35)

证：$A\vec{x} = P_{RanA}(\vec{b}) \in RanA \Rightarrow \vec{b} - A\vec{x} \in (RanA)^{\perp}$

$\Rightarrow \forall \vec{v} \in \mathbb{R}^n$，$A\vec{v} \in RanA$，则 $\langle \vec{b} - A\vec{x}, A\vec{v} \rangle = 0$

$\Rightarrow \forall \vec{v}$，$\langle A^T(\vec{b} - A\vec{x}), \vec{v} \rangle = 0$

$\Rightarrow A^T(\vec{b} - A\vec{x}) = 0 \Rightarrow A^TA\vec{x} = A^T\vec{b}$ 即正规方程. Q.E.D.

※ A满秩时，A^TA也满秩，正规方程 $A^TA\vec{x} = A^T\vec{b}$ 唯一最小二乘解

$\vec{x} = (A^TA)^{-1}A^T\vec{b}$ (7-36)

※ 另外 $\forall \vec{b} \in \mathbb{R}^n$，$P_{RanA}(\vec{b}) = A(A^TA)^{-1}A^T\vec{b}$

即 P_{RanA} 在 $\mathbb{R}^n$ 标准基下的矩阵为 $A(A^TA)^{-1}A^T$. (7-37)

· 线性回归

设 $\mathbb{R}^2$ 上有 n 个点 (x_i, y_i)，$i = 1, 2, \cdots, n$

为了拟合直线 $y = a + bx$，最小二乘法要求系数估计量 $\hat{a}$，$\hat{b}$ 使得

$\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 = \sum_{i=1}^{n}(\hat{a} + \hat{b}x_i - y_i)^2$ 取极小值.

即 $\|\hat{a} + \hat{b}\vec{x} - \vec{y}\|$ 取极小值，等价于求如下方程的最小二乘解.

$$\begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

设 $x_1, x_2, \cdots, x_n$ 不全为零，则左边的矩阵满秩.

$$\Rightarrow \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \left[\begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{pmatrix}\begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}\right]^{-1}\begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{pmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^{n} x_i^2 & -\sum_{i=1}^{n} x_i \\ -\sum_{i=1}^{n} x_i & n \end{pmatrix}\begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{pmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \cdot \frac{1}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

$$= \frac{1}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}\begin{pmatrix} \sum_{i=1}^{n}\sum_{j=1}^{n} x_i^2 y_j - \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j y_j \\ n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j \end{pmatrix}$$

即 $\hat{b} = \dfrac{\overline{xy} - \bar{x}\cdot\bar{y}}{\overline{x^2} - (\bar{x})^2}$, $\hat{a} = \dfrac{\overline{x^2}\cdot\bar{y} - \bar{x}\cdot\overline{xy}}{\overline{x^2} - (\bar{x})^2} = \bar{y} - \hat{b}\bar{x}$ (7-38)

※ 正规方程还可应用于各种线性/非线性拟合.

如二次函数: $y = a + bx + cx^2$

$\sum_{i=1}^{n} (\hat{a} + \hat{b}x_i + \hat{c}x_i^2 - y_i)^2$ 最小.

平面: $z = a + bx + cy$

$\sum_{i=1}^{n} (\hat{a} + \hat{b}x_i + \hat{c}y_i - z_i)^2$ 最小.

7.4 伴随与自伴算子

▸ 线性映射的伴随

• $\forall T \in \mathcal{L}(V, W)$, T的伴随 $T^*: W \to V$ 定义为:

$\forall \vec{v} \in V, \vec{w} \in W$. $\langle T(\vec{v}), \vec{w} \rangle = \langle \vec{v}, T^*(\vec{w}) \rangle$.

Riesz表示定理保证了存在性&唯一性:

$\forall T \in \mathcal{L}(V, W)$, 取定 $\vec{w} \in W$, 考虑 $\varphi \in V^*: \vec{v} \mapsto \langle T(\vec{v}), \vec{w} \rangle$.

由Riesz表示定理, 存在唯一的 $\vec{u} \in V$ 使得 $\varphi(\vec{v}) = \langle \vec{v}, \vec{u} \rangle$.

令 $\vec{u} = T^*(\vec{w})$, 得 T^* 是良定义的映射.

※ T^* 依赖于Riesz表示定理, 这要求 V 是有限维的.

为方便定义 $(T^*)^*$, W也取有限维向量空间. (本章剩余部分默认研究有限维

※ 这里定义的"伴随"与§3.3定义的伴随毫无关系!

• 伴随是线性映射: 若 $T \in \mathcal{L}(V, W)$, 则 $T^* \in \mathcal{L}(W, V)$ (7-39)

证: $\forall \vec{w}_1, \vec{w}_2 \in W$, 都有 $\forall \vec{v} \in V$,

$$\begin{aligned}\langle \vec{v}, T^*(\vec{w}_1 + \vec{w}_2) \rangle &= \langle T(\vec{v}), \vec{w}_1 + \vec{w}_2 \rangle \\ &= \langle T(\vec{v}), \vec{w}_1 \rangle + \langle T(\vec{v}), \vec{w}_2 \rangle \\ &= \langle \vec{v}, T^*(\vec{w}_1) \rangle + \langle \vec{v}, T^*(\vec{w}_2) \rangle \\ &= \langle \vec{v}, T^*(\vec{w}_1) + T^*(\vec{w}_2) \rangle\end{aligned}$$

$\Rightarrow T^*(\vec{w}_1 + \vec{w}_2) = T^*(\vec{w}_1) + T^*(\vec{w}_2)$

$\forall \vec{w} \in W, \lambda \in F$, 都有 $\forall \vec{v} \in V$.

$$
\begin{aligned}
<\vec{v}, T^*(\lambda \vec{w})> &= <T(\vec{v}), \lambda \vec{w}> \\
&= \bar{\lambda} <T(\vec{v}), \vec{w}> \\
&= \bar{\lambda} <\vec{v}, T^*(\vec{w})> \\
&= <\vec{v}, \lambda T^*(\vec{w})>
\end{aligned}
$$

$\Rightarrow T^*(\lambda \vec{w}) = \lambda T^*(\vec{w})$ Q.E.D.

- 伴随的基本性质 (7-40)

i) $\forall S, T \in \mathcal{L}(V, W)$, $(S+T)^* = S^* + T^*$

ii) $\forall \lambda \in \mathbb{F}$, $T \in \mathcal{L}(V, W)$, $(\lambda T)^* = \bar{\lambda} T^*$

iii) $\forall T \in \mathcal{L}(V, W)$, $(T^*)^* = T$

iv) $\forall S \in \mathcal{L}(W, U)$, $T \in \mathcal{L}(V, W)$, $(ST)^* = T^* S^*$

v) $I^* = I$

证: i) $\forall \vec{v} \in V, \vec{w} \in W$, $<\vec{v}, (S+T)^*(\vec{w})> = <(S+T)(\vec{v}), \vec{w}>$

$= <S(\vec{v}) + T(\vec{v}), \vec{w}> = <S(\vec{v}), \vec{w}> + <T(\vec{v}), \vec{w}>$

$= <\vec{v}, S^*(\vec{w})> + <\vec{v}, T^*(\vec{w})> = <\vec{v}, (S^* + T^*)(\vec{w})>$

$\Rightarrow (S+T)^* = S^* + T^*$

ii) $\forall \vec{v} \in V, \vec{w} \in W$, $<\vec{v}, (\lambda T)^*(\vec{w})> = <\lambda T(\vec{v}), \vec{w}>$

$= \lambda <T(\vec{v}), \vec{w}> = \lambda <\vec{v}, T^*(\vec{w})> = <\vec{v}, \bar{\lambda} T^*(\vec{w})>$

$\Rightarrow (\lambda T)^* = \bar{\lambda} T^*$

iii) $\forall \vec{v} \in V, \vec{w} \in W$, $<\vec{v}, (T^*)^*(\vec{w})> = <T^*(\vec{v}), \vec{w}>$

$= \overline{<\vec{w}, T^*(\vec{v})>} = \overline{<T(\vec{w}), \vec{v}>} = <\vec{v}, T(\vec{w})>$

$\Rightarrow (T^*)^* = T$

iv) $\forall \vec{v} \in V, \vec{u} \in U$, $<\vec{v}, (ST)^*(\vec{u})> = <(ST)(\vec{v}), \vec{u}>$

$= <T(\vec{v}), S^*(\vec{u})> = <\vec{v}, (T^* S^*)(\vec{u})>$

$\Rightarrow (ST)^* = T^* S^*$

v) $\forall \vec{v}, \vec{u} \in V$, $<\vec{v}, I^*\vec{u}> = <I\vec{v}, \vec{u}> = <\vec{v}, I\vec{u}>$.

$\Rightarrow I^* = I$ Q.E.D.

- T^*的基本子空间 (7-41)

i) $Null\,T^* = (Ran\,T)^\perp$

ii) $Ran\,T^* = (Null\,T)^\perp$

iii) $Null\,T = (Ran\,T^*)^\perp$

iv) $Ran\,T = (Null\,T^*)^\perp$

证: i) $\forall \vec{w} \in W$,

$\vec{w} \in Null\,T^* \Longleftrightarrow T^*(\vec{w}) = 0$

$\Longleftrightarrow <\vec{v}, T^*(\vec{w})> = 0$, $\forall \vec{v} \in V$

$\Longleftrightarrow <T(\vec{v}), \vec{w}> = 0$, $\forall \vec{v} \in V$

$\Longleftrightarrow \vec{w} \in (Ran\,T)^\perp$

$\Rightarrow Null\,T^* = (Ran\,T)^\perp$

iv) 对 i) 两端取正交补

ii) 把 iv) 的 T 与 T^* 对换, iii) 把 i) 的 T 与 T^* 对换. Q.E.D.

※ 这个结论揭示了矩阵四个基本子空间的重要联系.

▸ 矩阵的共轭转置

$\forall A \in \mathbb{F}^{m,n}$, A 的共轭转置 $A^* = \overline{A^T}$, 即 A 的转置的所有元素取复共轭.

※ 若 $A \in \mathbb{R}^{m,n}$, 显然 $A^* = A^T$, 使用 A^* 代替 A^T 有时避免分类讨论, 从而简化过程.

- 映射的伴随对应矩阵的共轭转置

设 $\vec{e_1}, \vec{e_2}, \cdots, \vec{e_n}$ 是 V 的标准正交基, $\vec{f_1}, \vec{f_2}, \cdots, \vec{f_m}$ 是 W 的标准正交基. 若 $T \in \mathcal{L}(V, W)$ 在该基下的矩阵为 A, 则 $T^* \in \mathcal{L}(W, V)$ 在该基下的矩阵为 A^*. (7-42)

证：设 T 在该基下矩阵为 $A=\{a_{ij}\}$，T^* 矩阵为 $B=\{b_{ij}\}$

$T(\vec{e_i})=\sum_{j=1}^{n}<T(\vec{e_i}),\vec{f_j}>\vec{f_j}=\sum_{j=1}^{m}a_{ij}\vec{f_j}$

$\Rightarrow a_{ij}=<T(\vec{e_i}),\vec{f_j}>$ 同理，对 T^*，$b_{ij}=<T^*(\vec{f_i}),\vec{e_j}>$

$\Rightarrow a_{ij}=<T(\vec{e_i}),\vec{f_j}>=<\vec{e_i},T^*(\vec{f_j})>=\overline{<T^*(\vec{f_j}),\vec{e_i}>}=\overline{b_{ji}}$，

$\forall i=1,2,\cdots m;\ j=1,2,\cdots,$

$\Rightarrow B=\overline{A^T}=A^*$ Q.E.D.

- 共轭转置的性质． (7-43)

i) $\forall A,B\in\mathbb{F}^{m,n}$，$(A+B)^*=A^*+B^*$

ii) $\forall A\in\mathbb{F}^{m,n}$，$\lambda\in\mathbb{F}$，$(\lambda A)^*=\bar{\lambda}A^*$

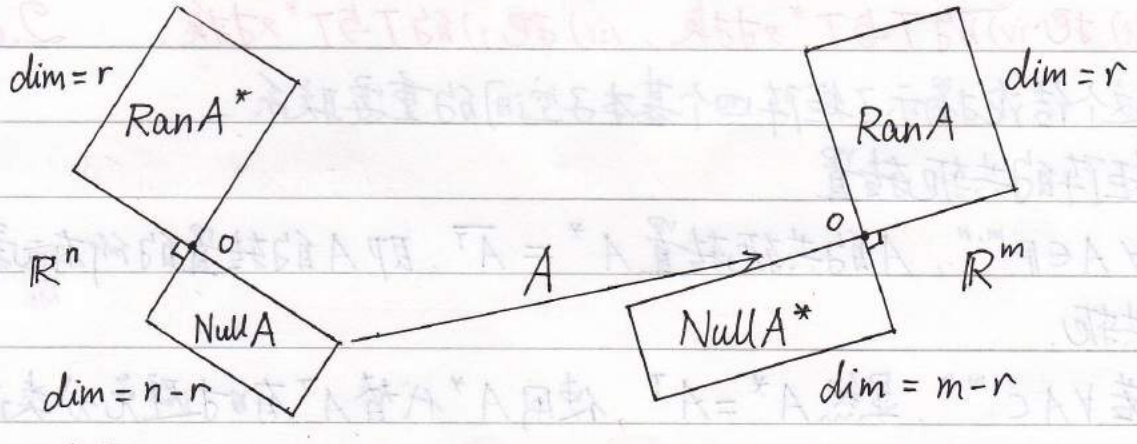
iii) $\forall A\in\mathbb{F}^{m,n}$，$B\in\mathbb{F}^{n,p}$，$(AB)^*=B^*A^*$

iv) $\forall A\in\mathbb{F}^{m,n}$，$(A^*)^*=A$

- 考虑实矩阵的4个基本子空间 $Null\,A$，$Null\,A^*$，$Ran\,A$，$Ran\,A^*$．

$\forall A\in\mathbb{R}^{m,n}$，有 $Null\,A=(Ran\,A^*)^\perp$，$Null\,A^*=(Ran\,A)^\perp$

$\Rightarrow \mathbb{R}^n=Null\,A\oplus Ran\,A^*$，$\mathbb{R}^m=Null\,A^*\oplus Ran\,A$ (7-44)



▸ 自伴算子

- $T\in\mathcal{L}(V)$ 称为自伴算子，若 $T=T^*$．

即 $\forall\vec{v},\vec{w}\in V$，$<T(\vec{v}),\vec{w}>=<\vec{v},T(\vec{w})>$．

※ $\mathcal{L}(V)$ 上的自伴算子 $T=T^*$ 可以类比 $\mathbb{C}$ 上的实数 $z=\bar{z}$

- 自伴算子的矩阵

设$T\in\mathcal{L}(V)$在V的一组标准正交基下成矩阵A. 若T是自伴算子，则有$A=A^*$，这样的矩阵称为Hermite矩阵. 实向量空间中的Hermite矩阵就是对称矩阵. ($A=A^T$) (7-45)

- 酉空间中的自伴算子有许多良好的性质：
- 自伴算子的本征值是实数. (7-46)

证：设$T\in\mathcal{L}(V)$自伴，若λ是T的本征值，则$\exists\vec{v}\neq 0$ s.t. $T(\vec{v})=\lambda\vec{v}$.

$\Rightarrow \lambda\|\vec{v}\|^2=\langle\lambda\vec{v},\vec{v}\rangle=\langle T(\vec{v}),\vec{v}\rangle=\langle\vec{v},T(\vec{v})\rangle=\langle\vec{v},\lambda\vec{v}\rangle=\bar{\lambda}\|\vec{v}\|^2$

$\vec{v}\neq 0\Rightarrow\lambda=\bar{\lambda}\Rightarrow\lambda\in\mathbb{R}$ Q.E.D.

- 设V是酉空间，$T\in\mathcal{L}(V)$自伴当且仅当$\forall\vec{v}\in V$，$\langle T(\vec{v}),\vec{v}\rangle\in\mathbb{R}$. (7-47)

证：$\langle T(\vec{v}),\vec{v}\rangle\in\mathbb{R}\iff\forall\vec{v}\in V$，$\langle T(\vec{v}),\vec{v}\rangle-\overline{\langle T(\vec{v}),\vec{v}\rangle}=0$

$\langle T(\vec{v}),\vec{v}\rangle-\overline{\langle T(\vec{v}),\vec{v}\rangle}=\langle T(\vec{v}),\vec{v}\rangle-\langle\vec{v},T(\vec{v})\rangle$

$=\langle(T-T^*)(\vec{v}),\vec{v}\rangle=0$

这里还不能直接认为$T-T^*=0$.

考虑Euclid空间旋转$\frac{\pi}{2}$的算子R，有$\forall\vec{v}\in V$，$\langle R(\vec{v}),\vec{v}\rangle=0$而$R\neq 0$.

酉空间的限制是必要的

$\forall\vec{u},\vec{w}\in V$，若$S\in\mathcal{L}(V)$使得$\forall\vec{v}\in V$，$\langle S(\vec{v}),\vec{v}\rangle=0$.

则有$\langle S(\vec{u}),\vec{w}\rangle=\frac{1}{4}[\langle S(\vec{u}+\vec{w}),\vec{u}+\vec{w}\rangle-\langle S(\vec{u}-\vec{w}),\vec{u}-\vec{w}\rangle$

$+i\langle S(\vec{u}+i\vec{w}),\vec{u}+i\vec{w}\rangle-i\langle S(\vec{u}-i\vec{w}),\vec{u}-i\vec{w}\rangle]$ (类似7-15)

$=0$ (每项都为零).

这时取$\vec{w}=S(\vec{u})\Rightarrow\forall\vec{u}\in V$，$\langle S(\vec{u}),S(\vec{u})\rangle=0\Rightarrow S=0$.

这样就证明了$T-T^*=0\Rightarrow T=T^*$. Q.E.D.

- 自伴算子不同本征值对应的本征向量正交. (7-48)

证：设$T\in\mathcal{L}(V)$，λ,μ是T的不同本征值，则$\exists\vec{v},\vec{w}\neq 0$，$T(\vec{v})=\lambda\vec{v}$，$T(\vec{w})=\mu\vec{w}$.

若$T=T^*$，由7-46，有$\lambda=\bar{\lambda}$，$\mu=\bar{\mu}$

$\lambda\langle\vec{v},\vec{w}\rangle=\langle\lambda\vec{v},\vec{w}\rangle=\langle T(\vec{v}),\vec{w}\rangle=\langle\vec{v},T(\vec{w})\rangle=\langle\vec{v},\mu\vec{w}\rangle$

$=\mu\langle\vec{v},\vec{w}\rangle$

$\lambda\neq\mu\Rightarrow\langle\vec{v},\vec{w}\rangle=0$，$\vec{v}$与$\vec{w}$正交 Q.E.D.

- 自伴算子的不变子空间 (7-49)

设$T\in\mathcal{L}(V)$自伴，若U是T的不变子空间，则：

i) $U^\perp$是T的不变子空间

ii) $T|_U\in\mathcal{L}(U)$自伴

iii) $T|_{U^\perp}\in\mathcal{L}(U^\perp)$自伴

证：i) $\forall\vec{v}\in U^\perp$，$\vec{u}\in U$，有$T(\vec{u})\in U$

$\langle T(\vec{v}),\vec{u}\rangle=\langle\vec{v},T(\vec{u})\rangle=0\Rightarrow T(\vec{v})\in U^\perp$

即$U^\perp$是T的不变子空间

ii) $\forall\ \vec{u},\vec{v}\in U$，

$\langle T|_U(\vec{u}),\vec{v}\rangle=\langle T(\vec{u}),\vec{v}\rangle=\langle\vec{u},T(\vec{v})\rangle=\langle\vec{u},T|_U(\vec{v})\rangle$

$\Rightarrow T|_U$是自伴算子

iii) 把ii)中的U换成$U^\perp$即得证. Q.E.D.

※ 定理7-48与7-49都可以弱化条件，推广到任意正规算子的情况，见§7.6关于正规算子的讨论.

7.5 实谱定理

- 在§5.2中，我们讨论了算子具有对角矩阵的各种条件，其中最重要的是本征向量张成一组基. 本节则讨论内积空间上具有对角矩阵的算子，它们的本征向量张成一组标准正交基. 即正交对角化问题.
- 谱定理描述了可正交对角化的矩阵：在Euclid空间中是自伴算子(对应对称矩阵)，在酉空间中是正规算子.

▸ 实谱定理

· Euclid空间上的自伴算子有本征值. (7-50)

证：自伴算子的本征值是实数 (7-46)，取标准正交基，对应的Hermite矩阵的特征多项式有实根.

Euclid空间中的自伴算子在标准正交基下成对称矩阵，其特征多项式的根全为实数，都是算子的本征值. 即Euclid空间上的算子有本征值.

Q.E.D.

· 实谱定理 (7-51)

设 T 是 Euclid 空间 V 上的算子，以下表述等价：

i) T 是自伴算子

ii) T 的本征向量在 V 中张成一组标准正交基

iii) 存在一组 V 上的标准正交基使 T 成对角矩阵.

证：ii) $\Leftrightarrow$ iii) 显然 (5-11)

iii) $\Rightarrow$ i)：对角矩阵是对称矩阵，那么 T 是自伴算子.

i) $\Rightarrow$ ii)：对 $\dim V$ 作归纳法：

a) $\dim V = 1$ 时，显然成立.

b) 假设 $\dim V = n-1$ 时，自伴算子的本征向量张成一组标准正交基

c) $\dim V = n$. 设 T 有一个本征值 λ_0 (7-50).

$\exists\ \vec{v}_0 \in V$ s.t. $T(\vec{v}_0) = \lambda_0 \vec{v}_0$，$\|\vec{v}_0\| = 1$.

设 $U = \mathrm{span}(\vec{v}_0)$ 在 T 下不变，则 $U^\perp$ 也在 T 下不变.

且 $T|_{U^\perp}$ 自伴. (7-49).

$\dim U^\perp = \dim V - \dim U = n-1$，按假设 b)，存在一组 $U^\perp$ 上的标准正交基 $\vec{v}_1, \cdots, \vec{v}_{n-1}$，由 $T|_{U^\perp}$ 的本征向量组成.

由于 $\|\vec{v}_0\| = 1$，$\vec{v}_0 \perp U^\perp$，则 $\vec{v}_0, \vec{v}_1, \cdots, \vec{v}_{n-1}$ 是 V 上的标准正交基，且都是 T 的本征向量.

Q.E.D.

※ 实对称矩阵可相似对角化，现在再使之正交对角化：

· 正交对角化

若A是实对称矩阵，则存在正交矩阵P，使得$A=PDP^{-1}$，其中D是对角矩阵. (7-52)

(正交矩阵各列构成一组列空间的标准正交基，详见7.7)

证：把A看作V内一个自伴算子T在标准正交基$\vec{e_1},\vec{e_2},\cdots,\vec{e_n}$下的矩阵. 按实谱定理，$V$上存在另一组标准正交基$\vec{f_1},\vec{f_2},\cdots,\vec{f_n}$使$T$成对角形$D$.

则A与D相似，且过渡矩阵P是正交矩阵

$\Rightarrow A=PDP^{-1}$. Q.E.D.

· 实二次型的正交换元

设实二次型$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\ (a_{ij}=a_{ji})$，存在正交矩阵$P\in\mathbb{R}^{n,n}$，使二次型在正交换元$\vec{x}=P\vec{z}$下化成标准形：$f=\sum_{i=1}^{n}\lambda_iz_i^2$.

其中$\lambda_1,\lambda_2,\cdots,\lambda_n$是$f$对应的矩阵$A$的全部本征值. (7-53)

证：f的矩阵A是实对称矩阵，由7-52，存在正交矩阵P，使

$A=PDP^{-1}$，$D=diag(\lambda_1,\lambda_2,\cdots,\lambda_n)$.

$$f=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j=\vec{x}^TA\vec{x}\xlongequal{\vec{x}=P\vec{z}}(P\vec{z})^TA(P\vec{z})$$

$$=\vec{z}^T(P^TAP)\vec{z}=\vec{z}^T(P^{-1}AP)\vec{z}=\vec{z}^TD\vec{z}$$

$$=\sum_{i=1}^{n}\lambda_iz_i^2$$ (正交矩阵$P^T=P^{-1}$，详见7.7) Q.E.D.

▸ 正交对角化算法

· 设$E(\lambda_i,T)$是T的本征值λ_i对应的本征空间，且$V=\bigoplus_{i=1}^{m}E(\lambda_i,T)$，则每个$E(\lambda_i,T)$取一组标准正交基，可以并成$V$的一组标准正交基. (7-54)

证：由定理4-36与7-48，易证.

~~※ 以下算法使用了V与$\mathbb{R}^n$等距同构这一事实，详见7.7~~

· 求正交矩阵P与对角矩阵D，使实对称矩阵$A=PDP^{-1}$. (7-55)

i) 计算特征多项式$det(A-\lambda I)$的根$\lambda_1,\lambda_2,\cdots,\lambda_m$

ii) 对每个λ_i，求出一组$Null(A-\lambda_iI)$的基

iii) 将求出的基用Gram-Schmidt过程得到$Null(A-\lambda_i I)$的标准正交基

iv) 每个$Null(A-\lambda_i I)$的标准正交基并成$\mathbb{R}^n$的基，每个基向量作为矩阵的列并成过渡矩阵P

v) 将本征值$\lambda_1,\cdots,\lambda_m$依次排列在D的对角线上，其中每个$\lambda_i$重复$\dim Null(A-\lambda_i I)$次.

- 求二次型 $f=\sum_{i=1}^n\sum_{j=1}^n a_{ij}x_ix_j$ 的标准形： (7-56)

i) 写出对应的矩阵A

ii) 求出正交矩阵P，使A正交对角化. $A=PDP^{-1}$

iii) 换元 $\vec{x}=P\vec{z}$ 化二次型为标准形 $f=\sum_{i=1}^n\lambda_i z_i^2$

7.6 正规算子与复谱定理

▶ 正规算子

- $T\in\mathcal{L}(V)$ 称为正规算子，若 $TT^*=T^*T$

※ 自伴算子显然也是正规算子.

- $T\in\mathcal{L}(V)$ 正规当且仅当 $\forall\vec{v}\in V$, $\|T(\vec{v})\|=\|T^*(\vec{v})\|$ (7-57)

证：必要性：若$T\in\mathcal{L}(V)$正规，$TT^*=T^*T$

$\Rightarrow\forall\vec{v}\in V$, $\langle TT^*(\vec{v}),\vec{v}\rangle=\langle T^*T(\vec{v}),\vec{v}\rangle$

$\Rightarrow\forall\vec{v}\in V$, $\langle T^*(\vec{v}),T^*(\vec{v})\rangle=\langle T(\vec{v}),T(\vec{v})\rangle$

$\Rightarrow\forall\vec{v}\in V$, $\|T(\vec{v})\|=\|T^*(\vec{v})\|$

充分性：若$\forall\vec{v}\in V$, $\|T(\vec{v})\|=\|T^*(\vec{v})\|$

$\Rightarrow\forall\vec{v}\in V$, $\langle TT^*(\vec{v}),\vec{v}\rangle=\langle T^*T(\vec{v}),\vec{v}\rangle$

令 $S=TT^*-T^*T$，于是有 $\forall\vec{v}\in V$, $\langle S(\vec{v}),\vec{v}\rangle=0$ 且S是自伴算子

参考7-47的证明，有 $S=0$

$\Rightarrow TT^*=T^*T$，T是正规算子 Q.E.D.

- 设$T\in\mathcal{L}(V)$，$\lambda\in\mathbb{F}$是T的本征值当且仅当$\bar{\lambda}\in\mathbb{F}$是$T^*$的本征值. (7-58)

证：λ是T的本征值 $\Longleftrightarrow \exists \vec{v}\neq 0$，$T(\vec{v})=\lambda\vec{v}$.

$\Longleftrightarrow \exists \vec{v}\neq 0, \forall \vec{w}\in V, \langle\vec{v},T^*(\vec{w})\rangle=\langle T(\vec{v}),\vec{w}\rangle=\langle\lambda\vec{v},\vec{w}\rangle$

$=\langle\vec{v},\bar{\lambda}\vec{w}\rangle$

$\Longleftrightarrow \exists \vec{v}\neq 0, \forall \vec{w}\in V, \langle\vec{v},(T^*-\bar{\lambda}I)(\vec{w})\rangle=0$.

$\Longleftrightarrow \exists \vec{v}\neq 0, \vec{v}\perp Ran(T^*-\bar{\lambda}I)$.

$\Longleftrightarrow T^*-\bar{\lambda}I$不可逆

$\Longleftrightarrow Null(T^*-\bar{\lambda}I)\neq\{0\}$.

$\Longleftrightarrow \bar{\lambda}$是$T^*$的本征值. Q.E.D.

※ 当T是正规算子时，该定理有加强版：

- 设$T\in\mathcal{L}(V)$正规，$\vec{v}\in V$是T对应本征值λ的本征向量，则$\vec{v}$也是T^*对应本征值$\bar{\lambda}$的本征向量. (7-59)

证：$TT^*=T^*T \Rightarrow (T-\lambda I)(T-\lambda I)^*=(T-\lambda I)(T^*-\bar{\lambda}I)$

$=TT^*-\lambda T^*-\bar{\lambda}T+|\lambda|^2 I=T^*T-\bar{\lambda}T-\lambda T^*+|\lambda|^2 I$

$=(T^*-\bar{\lambda}I)(T-\lambda I)=(T-\lambda I)^*(T-\lambda I)$

$\Rightarrow T-\lambda I$也是正规算子

$\Rightarrow \forall \vec{w}\in V, \|(T-\lambda I)(\vec{w})\|=\|(T^*-\bar{\lambda}I)(\vec{w})\|$

$\vec{v}$是本征向量，$0=\|(T-\lambda I)(\vec{v})\|=\|(T^*-\bar{\lambda}I)(\vec{v})\|$.

这说明$\vec{v}$也是T^*对应$\bar{\lambda}$的本征向量. Q.E.D.

- 推广7-48：正规算子不同本征值对应的本征值正交. (7-60)

证：设$T\in\mathcal{L}(V)$正规，λ,μ是T的不同本征值. $\exists \vec{v},\vec{w}\neq 0$, $T(\vec{v})=\lambda\vec{v}$, $T(\vec{w})=\mu\vec{w}$. 由7-59，有$T^*(\vec{w})=\bar{\mu}\vec{w}$.

$\lambda\langle\vec{v},\vec{w}\rangle=\langle T(\vec{v}),\vec{w}\rangle=\langle\vec{v},T^*(\vec{w})\rangle=\langle\vec{v},\bar{\mu}\vec{w}\rangle=\mu\langle\vec{v},\vec{w}\rangle$.

$\lambda\neq\mu\Rightarrow\langle\vec{v},\vec{w}\rangle=0$，$\vec{v}$与$\vec{w}$正交. Q.E.D.

- 推广7-49：正规算子的不变子空间

设$T\in\mathcal{L}(V)$正规，若U是T的不变子空间，则： (7-61)

i) $U^{\perp}$是T的不变子空间

ii) U是T^*的不变子空间

iii) $(T|_U)^* = T^*|_U$

iv) $T|_U \in \mathcal{L}(U)$与$T^*|_{U^\perp} \in \mathcal{L}(U^\perp)$正规

证：i) 设U有标准正交基$\vec{e_1}, \vec{e_2}, \cdots, \vec{e_m}$，$T|_U$在此基下矩阵为$A = \{a_{i,j}\}$.

将U的标准正交基扩充为V的标准正交基$\vec{e_1}, \cdots, \vec{e_m}, \vec{f_1}, \cdots, \vec{f_n}$ (7-34).

设T在此基下的矩阵$M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$，其中$B = \{b_{i,j}\}$.

$$\sum_{j=1}^{m} \|T(\vec{e_j})\|^2 = \sum_{j=1}^{m} \|\sum_{i=1}^{m} a_{i,j} \vec{e_i}\|^2 = \sum_{i=1}^{m}\sum_{j=1}^{m} |a_{i,j}|^2.$$

T^*在此基下的矩阵$M^* = \begin{pmatrix} A^* & 0 \\ B^* & C^* \end{pmatrix}$，$A^* = \{\overline{a_{j,i}}\}$，$B^* = \{\overline{b_{j,i}}\}$.

$$\sum_{j=1}^{m} \|T^*(\vec{e_j})\|^2 = \sum_{j=1}^{m} \|\sum_{i=1}^{m} \overline{a_{j,i}}\vec{e_i} + \sum_{k=1}^{n} \overline{b_{j,k}}\vec{e_k}\|^2 = \sum_{j=1}^{m} \left(\sum_{i=1}^{m} |a_{j,i}|^2 + \sum_{k=1}^{m} |b_{j,k}|^2\right)$$
$$= \sum_{i=1}^{m}\sum_{j=1}^{m} |a_{i,j}|^2 + \sum_{j=1}^{m}\sum_{k=1}^{n} |b_{k,j}|^2$$

由7-57，$\forall j = 1, \cdots, m$，$\|T(\vec{e_j})\| = \|T^*(\vec{e_j})\|$

$\Rightarrow \sum_{j=1}^{m} \|T(\vec{e_j})\|^2 = \sum_{j=1}^{m} \|T^*(\vec{e_j})\| \Rightarrow \sum_{j=1}^{m}\sum_{k=1}^{n} |b_{k,j}|^2 = 0 \Rightarrow B = 0$.

$\Rightarrow M = \begin{pmatrix} A & 0 \\ 0 & C \end{pmatrix}$，即$U^{\perp} = \mathrm{span}(\vec{f_1}, \cdots, \vec{f_n})$是T的不变子空间.

ii) $M^* = \begin{pmatrix} A^* & 0 \\ 0 & C^* \end{pmatrix}$，即$U = \mathrm{span}(\vec{e_1}, \cdots, \vec{e_m})$是$T^*$的不变子空间.

iii) 设$S = T|_U$. $\forall \vec{u}, \vec{v} \in U$,

$\langle S(\vec{u}), \vec{v} \rangle = \langle T(\vec{u}), \vec{v} \rangle = \langle \vec{u}, T^*(\vec{v}) \rangle$

$\Rightarrow S^*(\vec{v}) = T^*(\vec{v})$, $\forall \vec{v} \in U \Rightarrow T^*|_U = S^* = (T|_U)^*$

iv) $TT^* = T^*T \Rightarrow T|_U T^*|_U = T^*|_U T|_U$.

$\Rightarrow T|_U (T|_U)^* = (T|_U)^* T|_U \Rightarrow T|_U$正规

交换U与$U^{\perp}$即得$T^*|_{U^\perp}$也正规. Q.E.D.

▸ 复谱定理

· 实谱定理描述了Euclid空间中自伴算子可正交对角化. 此结论可以直接移植到酉空间上. 即: 酉空间上有一组由自伴算子的本征向量构成的标准正交基, 使算子在该基下成实对角矩阵. (7-62)

· 与Euclid空间不同的是, 代数基本定理保证酉空间上的算子天然就有足够多的本征值. 因此酉空间上可正交对角化的不只有自伴算子. 下面将证明酉空间有且只有正规算子可正交对角化.

· 复谱定理

设T是酉空间V上的算子, 以下表述等价: (7-63)

i) T是正规算子

ii) T的本征向量在V中构成一组标准正交基

iii) 存在一组V上的标准正交基使T成对角矩阵.

证一: 在实谱定理的证明中用到了7-48 & 7-49. 而对正规算子也有几乎一样的7-60 & 7-61. 复谱定理的证明与实谱定理也几乎一样了.

ii) <=> iii) 显然

iii) => i) 对角矩阵显然与共轭转置可交换, 那么T是正规算子

i) ⇒ ii) 对dim V作归纳法

a) dim V = 1 时显然成立

b) 假设 dim V = n-1 时, 正规算子本征向量构成标准正交基

c) 若dim V = n. 设 λ_0 是T的本征值

$\exists \vec{v_0} \in V$, $T(\vec{v_0}) = \lambda_0 \vec{v_0}$, $\|\vec{v_0}\| = 1$.

$U = span(\vec{v_0})$ 在T下不变, 则 $U^\perp$ 也在T下不变, 且 $T|_{U^\perp}$ 正规. (7-61)

$\dim U^\perp = \dim V - \dim U = n-1$, 按假设b), $U^\perp$ 上有 $T|_{U^\perp}$ 本征向量构成的标准正交基 $\vec{v_1}, \vec{v_2}, \cdots, \vec{v_{n-1}}$.

由于 $\|\vec{v_0}\| = 1$, $\vec{v_0} \perp U^\perp$, 于是 $\vec{v_0}, \vec{v_1}, \cdots, \vec{v_{n-1}}$ 是V上标准正交基. Q.E.D.

证二：从另一种思路出发也能得到复谱定理，即利用已证明的Schur定理。

i) ⇒ iii) 由Schur定理，设T在V的标准正交基$\vec{e_1},\vec{e_2},\cdots,\vec{e_n}$下成上三角矩阵

$$A=\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ & \ddots & \vdots \\ 0 & & a_{n,n} \end{pmatrix}$$

下面证明A其实是个对角矩阵～

$$T^*\text{在该基下矩阵为}A^*=\begin{pmatrix} \overline{a_{1,1}} & & 0 \\ \vdots & \ddots & \\ \overline{a_{1,n}} & \cdots & \overline{a_{n,n}} \end{pmatrix}$$

$\|T(\vec{e_1})\|^2=\|a_{1,1}\vec{e_1}\|^2=|a_{1,1}|^2$

$\|T^*(\vec{e_1})\|^2=\|\sum_{i=1}^{n}\overline{a_{1,i}}\,\vec{e_i}\|^2=\sum_{i=1}^{n}|a_{1,i}|^2$

T正规 $\Rightarrow \|T(\vec{e_1})\|=\|T^*(\vec{e_1})\| \Rightarrow |a_{1,1}|^2=\sum_{i=1}^{n}|a_{1,i}|^2$

$\Rightarrow a_{1,2}=a_{1,3}=\cdots=a_{1,n}=0$.（A第1行除$a_{1,1}$全为0）

$\|T(\vec{e_2})\|^2=|a_{1,2}|^2+|a_{2,2}|^2=|a_{2,2}|^2$

$\|T^*(\vec{e_2})\|^2=\sum_{i=2}^{n}|a_{2,i}|^2$

$\|T(\vec{e_2})\|=\|T^*(\vec{e_2})\| \Rightarrow a_{2,3}=a_{2,4}=\cdots=a_{2,n}=0$

以此类推，所以A所有非对角元全为0. A为对角矩阵. Q.E.D.

- Euclid空间中的正规算子

- 复谱定理给出了酉空间上正规算子的结构的完整刻画. 在Euclid空间中，虽然正规算子无法正交对角化（除非自伴），但它们在标准正交基上也能有比较简单的准对角形（由1阶与2阶方阵组成）.

  ~~其本质是$\mathbb{R}$上没有高于2次的既约多项式~~ ?

- 二维Euclid空间中的非自伴的正规算子结构.

  设T是Euclid空间V上算子，$\dim V=2$，以下表述等价： (7-64)

  i) T是非自伴的正规算子

ii) T关于V上的某个标准正交基有矩阵 $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$，其中 $b>0$.

iii) T关于V上的所有标准正交基都有矩阵 $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$，其中 $b\neq 0$

证：iii) ⇒ ii) 显然.

ii) ⇒ i) 设 $A=\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$，$b>0$.

$$AA^*=\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} a & b \\ -b & a \end{pmatrix}=\begin{pmatrix} a^2+b^2 & 0 \\ 0 & a^2+b^2 \end{pmatrix}=\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} a & -b \\ b & a \end{pmatrix}=A^*A$$

⇒ T是正规算子. $b\neq -b$，T不是自伴算子

i) ⇒ iii) 设T关于V上的一个标准正交基有矩阵：

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix},\ a,b,c,d\in \mathbb{R}$$

T是正规算子 ⇒ $AA^*=A^*A$

$$\Rightarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix}=\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix}=\begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{pmatrix}$$

$$\Rightarrow \begin{cases} b^2=c^2 \\ ac+bd=ab+cd \end{cases} \Rightarrow (a-d)(c-b)=0.$$

T非自伴，$A\neq A^*$ ⇒ $b\neq c$ ⇒ $c=-b$ ⇒ $d=a$.

$$\Rightarrow A=\begin{pmatrix} a & -b \\ b & a \end{pmatrix},\ b\neq 0$$

Q.E.D.

※ §5.8的5-64证明了实向量空间的算子都有一维或二维不变子空间. 利用这个结论可以把7-64关于二维实正规算子的刻画推广到更高维度上.

- 设 T 是 Euclid 空间 V 上的正规算子，则 V 有标准正交基使 T 在该基下成准对角形 $diag(\lambda_1, \cdots, \lambda_m, S_1, \cdots, S_k)$
  其中 $\lambda_1, \cdots, \lambda_m$ 是 T 的本征值；$S_1, \cdots, S_k$ 是 2 阶子方阵，
  $$S_i = \begin{pmatrix} a_i & -b_i \\ b_i & a_i \end{pmatrix}, \; b_i > 0.$$
  (7-65

证：对 $\dim V$ 作归纳法：
$\dim V = 1$，显然成立。
$\dim V = 2$，T 自伴的情况见实谱定理，非自伴的情况见 7-64。
假设结论对维数小于 V 的 Euclid 空间成立。
根据 T 是否有本征值分情况定义 U：
i) 若 T 有本征向量 $\vec{v}$，令 $U = span(\vec{v})$。
令 $\vec{u} = \frac{\vec{v}}{\|\vec{v}\|}$，则 $\vec{u}$ 是 U 的标准正交基。
ii) 若 T 没有本征向量，令 U 是 T 的二维不变子空间。
$T|_U$ 是 U 上的正规算子（定理 7-61），可以取 U 上的标准正交基使 $T|_U$ 成 S_i 的形式（定理 7-64）。
$V = U \oplus U^\perp$，$U^\perp$ 也是 T 的不变子空间，$T|_{U^\perp}$ 是 $U^\perp$ 上正规算子，
并有 $\dim U^\perp < \dim V$，对 $U^\perp$ 应用归纳假设：
$U^\perp$ 有标准正交基使 $T|_{U^\perp}$ 成所求准对角形；把 U 与 $U^\perp$ 的基并成 V 的标准正交基，T 在该基下成所求准对角形。 Q.E.D.

7.7 等距同构

- 等距同构

- 保持范数的线性映射同时保持内积。
  设 V, W 是定义在同一域 F 上的内积空间，若 $T \in \mathcal{L}(V, W)$ 使 $\forall \vec{v} \in V$，$\|T(\vec{v})\| = \|\vec{v}\|$，则 $\forall \vec{u}, \vec{v} \in V$，$\langle T(\vec{u}), T(\vec{v}) \rangle = \langle \vec{u}, \vec{v} \rangle$ (7-66)

证：i) 若 V, W 是 Euclid 空间，利用等式 7-14，

$$\langle T(\vec{u}), T(\vec{v})\rangle = \frac{1}{4}(\|T(\vec{u})+T(\vec{v})\|^2 - \|T(\vec{u})-T(\vec{v})\|^2)$$
$$= \frac{1}{4}(\|\vec{u}+\vec{v}\|^2 - \|\vec{u}-\vec{v}\|^2)$$
$$= \langle \vec{u}, \vec{v}\rangle$$

ii) 若 V, W 是酉空间，利用等式 7-16，

$$\langle T(\vec{u}), T(\vec{v})\rangle = \frac{1}{4}\sum_{n=1}^{4} i^n \|T(\vec{u})+i^n T(\vec{v})\|^2$$
$$= \frac{1}{4}\sum_{n=1}^{4} i^n \|T(\vec{u}+i^n \vec{v})\|^2$$
$$= \frac{1}{4}\sum_{n=1}^{4} i^n \|\vec{u}+i^n \vec{v}\|^2$$
$$= \langle \vec{u}, \vec{v}\rangle$$

Q.E.D.

- 保持范数的可逆线性映射称为等距同构。
- 等距同构的性质

设 $T\in\mathcal{L}(V,W)$，$\dim V = \dim W$，下列表述等价： (7-67)

i) T 保持范数不变

ii) T 保持内积不变

iii) T 是等距同构

iv) T 把 V 上的某个标准正交基映到 W 的标准正交基

v) T 把 V 上的每个标准正交基都映到 W 的标准正交基。

vi) $T^*T = I_V$

vii) $TT^* = I_W$

viii) $T^{-1} = T^*$

证：轮转证法：$i) \Rightarrow ii) \Rightarrow v) \Rightarrow iv) \Rightarrow vi) \Rightarrow viii) \Rightarrow i) \Longleftrightarrow iii)$，$vi) \Rightarrow vii) \Rightarrow viii)$

i) ⇒ ii)：在 7-66 中已证

ii) ⇒ v)：设 $\vec{e}_1, \cdots, \vec{e}_n$ 是 V 的标准正交基。

$$\langle T(\vec{e}_i), T(\vec{e}_j)\rangle = \langle \vec{e}_i, \vec{e}_j\rangle = \delta_{i,j}$$

$\Rightarrow T(\vec{e}_1), \cdots, T(\vec{e}_n)$ 是 W 的标准正交基。

v) ⇒ iv)：显然

iv) ⇒ vi)：设 V 有标准正交基 $\vec{e_1},\cdots,\vec{e_n}$ 使得 $T(\vec{e_1}),\cdots,T(\vec{e_n})$ 是 W 的标准正交基

$\Rightarrow \langle\vec{e_i},\vec{e_j}\rangle = \langle T(\vec{e_i}),T(\vec{e_j})\rangle$

$\forall \vec{u},\vec{v}\in V$，设 $\vec{u}=\sum_{i=1}^{n}\lambda_i\vec{e_i}$，$\vec{v}=\sum_{j=1}^{n}\varepsilon_j\vec{e_j}$

$$\langle T^*T(\vec{u}),\vec{v}\rangle = \langle T(\vec{u}),T(\vec{v})\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\varepsilon_j\langle T(\vec{e_i}),T(\vec{e_j})\rangle$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\varepsilon_j\langle\vec{e_i},\vec{e_j}\rangle = \langle\vec{u},\vec{v}\rangle$$

$\Rightarrow T^*T = I_V$

vi) ⇒ viii)：$T^*T = I_V$，由线性映射基本定理：

$\dim V = \dim \mathrm{Null}\, T + \dim \mathrm{Ran}\, T \geq \dim \mathrm{Ran}\, T$

$\dim \mathrm{Ran}\, T = \dim \mathrm{Null}\, T^* + \dim V \geq \dim V$

$\Rightarrow \dim \mathrm{Ran}\, T = \dim V = \dim W \Rightarrow \mathrm{Ran}\, T = W$，$\mathrm{Null}\, T = \{0\}$.

于是 T 可逆，$T^{-1}=T^*$，$TT^*=I_W$

viii) ⇒ i)：$T^{-1}=T^* \Rightarrow T^*T = I_V$.

$\forall \vec{v}\in V$，$\|T(\vec{v})\|^2 = \langle T(\vec{v}),T(\vec{v})\rangle = \langle T^*T(\vec{v}),\vec{v}\rangle$

$= \langle\vec{v},\vec{v}\rangle = \|\vec{v}\|^2$

⇒ T 保持范数

vii) ⇒ viii)：与 vi) ⇒ viii) 同理，可推出 $(T^*)^{-1}=T \Rightarrow T^{-1}=T^*$.

i) + viii) ⇒ iii)：定义

iii) ⇒ i)：显然.　　　　Q.E.D.

（这里有点乱…）

· 内积空间 V 与 W 是等距同构的，若存在 V 到 W 的一个等距同构.

· 有限维内积空间同构当且仅当维数相等.　　(7-68)

证：必要性显然.

充分性：设 $\dim V = \dim W$，V 上取标准正交基 $\vec{e_1},\cdots,\vec{e_n}$，$W$ 上取标准正交基 $\vec{f_1},\cdots,\vec{f_n}$. 设 $T: V\to W$ 使 $T(\vec{e_i})=\vec{f_i}$，$i=1,\cdots,n$.

则 $<T(\vec{e_i}),T(\vec{e_j})> = <\vec{f_i},\vec{f_j}> = \delta_{ij} = <\vec{e_i},\vec{e_j}>$

$\Rightarrow \|T(\vec{e_i})\| = \|\vec{e_i}\| \Rightarrow \forall \vec{v} \in V, \|T(\vec{v})\| = \|\vec{v}\|$

易证T是同构映射，上式说明T是等距同构，即V与W等距同构. Q.E.D.

※ 关于等距同构的算子，不同的教材给出了两套说法(名词)：

A. 若等距同构 $S \in \mathcal{L}(V)$，则 S 又称为酉算子

B. 若 V 是 Euclid 空间，等距同构 $S \in \mathcal{L}(V)$ 称为正交算子；

若 V 是酉空间，等距同构 $S \in \mathcal{L}(V)$ 称为酉算子.

习惯上个人倾向于B.

· 正交算子、酉算子都是正规算子，因为 $SS^* = S^*S = I$. (7-69)

▸ 等距同构的矩阵

· 正交矩阵/酉矩阵是 Euclid 空间/酉空间中标准正交基之间的过渡矩阵.

· 正交算子/酉算子在标准正交基下的矩阵是正交矩阵/酉矩阵. (7-70)

证：7-67的性质v)

· 正交矩阵/酉矩阵的性质： (7-71)

若A, B是正交矩阵/酉矩阵，则：

i) A可逆，$A^{-1} = A^*$

ii) A^{-1}是正交矩阵/酉矩阵

iii) AB是正交矩阵/酉矩阵.

· 正交相似/酉相似

若存在正交矩阵/酉矩阵U使得 $A = U^*BU = U^{-1}BU$，则称矩阵A与B正交相似/酉相似

※ 易验证正交相似/酉相似矩阵构成等价类.

※ 正交相似/酉相似的矩阵本质上是算子在不同标准正交基下的矩阵.

▸ 等距同构的行列式与本征值

· 设 $S \in \mathcal{L}(V)$ 是等距同构，则 $|\det S| = 1$. (7-72)

证：设 S 在某个标准正交基下有矩阵 U，$U^{-1}=U^*$，U 是酉矩阵.

设 U 的本征值为 $\lambda_1,\cdots,\lambda_n$（计重数）

$\Rightarrow U^*$ 的本征值为 $\overline{\lambda_1},\cdots,\overline{\lambda_n}$（定理 7-58）.

$\Rightarrow \det U^*=\overline{\lambda_1}\cdot\ldots\cdot\overline{\lambda_n}=\overline{\lambda_1\cdot\ldots\cdot\lambda_n}=\overline{\det U}$

$\Rightarrow \det U^*\cdot\det U=|\det U|^2$

又有 $\det U\cdot\det U^*=\det(UU^*)=\det I=1$

$\Rightarrow |\det U|^2=1 \Rightarrow |\det U|=1 \Rightarrow |\det S|=1$ Q.E.D.

※ 对正交算子 $S\in\mathcal{L}(V)$，$\det S=\pm1$. (7-73)

$\det S=1$ 的正交算子又称为旋转.

· 设 $S\in\mathcal{L}(V)$ 是等距同构，若 λ 是 S 的本征值，则 $|\lambda|=1$ (7-74)

证：$\exists\vec{v}\neq0$ s.t. $S(\vec{v})=\lambda\vec{v}$.

$\Rightarrow \|\vec{v}\|=\|S(\vec{v})\|=\|\lambda\vec{v}\|=|\lambda|\cdot\|\vec{v}\|$

$\Rightarrow |\lambda|=1$

▸ 酉算子的刻画

· 酉算子是酉空间上的正规算子，因此可以应用复谱定理：

V 中有一个由酉算子 S 的本征向量构成的标准正交基，S 在该基下成对角形，且对角元（本征值）的绝对值等于 1. (7-75)

▸ 正交算子的刻画

· 正交算子是 Euclid 空间上的正规算子，可以应用 7-65：

· 设 S 是 Euclid 空间上的正交算子，则 V 有标准正交基使 S 在该基下成准对角形 $\mathrm{diag}(\lambda_1,\cdots,\lambda_m,S_1,\cdots,S_k)$

其中 $\lambda_1,\cdots,\lambda_m$ 是 S 的本征值，$\lambda_i=\pm1$；$S_1,\cdots,S_k$ 是 2 阶方阵，

$$S_i=\begin{pmatrix}\cos\theta_i & -\sin\theta_i\\ \sin\theta_i & \cos\theta_i\end{pmatrix},\ \theta_i\in(0,\pi).$$ (7-76)

证：由 7-65，我们已知 S 在标准正交基成 $\mathrm{diag}(\lambda_1,\cdots,\lambda_m,S_1,\cdots,S_k)$

其中 $S_i=\begin{pmatrix} a_i & -b_i \\ b_i & a_i \end{pmatrix}$，$b_i>0$

i) 7-74得 $|\lambda_i|=1$，$\lambda_i\in\mathbb{R} \Rightarrow \lambda_i=\pm1$

ii) 对 S_i，∃基向量 $\vec{e_j}, \vec{e_{j+1}}$ s.t. $S(\vec{e_j})=a_i\vec{e_j}+b_i\vec{e_{j+1}}$

$\Rightarrow 1=\|\vec{e_j}\|^2=\|S(\vec{e_j})\|^2=\|a_i\vec{e_j}+b_i\vec{e_{j+1}}\|^2=a_i^2+b_i^2$

$\Rightarrow \exists\,\theta_i\in(0,\pi)$ s.t. $a_i=\cos\theta_i$，$b_i=\sin\theta_i$

$$\Rightarrow S_i=\begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix},\ \theta_i\in(0,\pi).$$

Q.E.D.

例： $\mathbb{R}^3$ 中的旋转矩阵（内积默认取点积）

设 S 是 $\mathbb{R}^3$ 中的一个旋转，即 S 是正交算子且 $\det S=1$.

由7-76，$\mathbb{R}^3$ 上有一组标准正交基 $\vec{e_1},\vec{e_2},\vec{e_3}$，使 S 的矩阵为

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix},\ \theta\in[0,\pi].$$

翻译成物理语言：

刚体的任意定点转动都等价于刚体绕瞬时转轴（这里的矢量 $\vec{e_1}$）作定轴转动；刚体上每点都有相同的角速度 $\vec{\omega}$.

- 矩阵的QR分解.

- 初等旋转（Givens旋转）

几何上初等旋转相当于Euclid空间中的"定轴平面旋转".

设 $\vec{e_1},\vec{e_2},\cdots,\vec{e_n}$ 是Euclid空间 V 的标准正交基，初等旋转 R：

$R(\vec{e_i})=\cos\theta\cdot\vec{e_i}+\sin\theta\cdot\vec{e_j}$

$R(\vec{e_j})=-\sin\theta\cdot\vec{e_i}+\cos\theta\cdot\vec{e_j}$

$R(\vec{e_k})=\vec{e_k}$，$\forall k\neq i,j$.

- 设 V 是Euclid空间，$\vec{e_1},\vec{e_2},\cdots,\vec{e_n}$ 是标准正交基，$\forall\vec{v}\in V$，∃ $n-1$ 个初等矩阵

$R_1, R_2, \cdots, R_{n-1}$ s.t. $R_{n-1} \cdots R_1(\vec{v}) = \|\vec{v}\| \vec{e}_1$ (7-77)

证：令 $R_i(\vec{e_1}) = \cos\theta_i \cdot \vec{e_1} + \sin\theta_i \cdot \vec{e_{i+1}}$

$R_i(\vec{e_{i+1}}) = -\sin\theta_i \cdot \vec{e_1} + \cos\theta_i \cdot \vec{e_{i+1}}$, $\theta_i = -\arctan\frac{\langle \vec{v_i}, \vec{e_{i+1}} \rangle}{\langle \vec{v_i}, \vec{e_1} \rangle}$

其中 $\vec{v_1} = \vec{v}$, $\vec{v_i} = R_{i-1}(\vec{v_{i-1}})$, $\forall i = 2, 3, \cdots, n-1$.

$\Rightarrow R_i(\vec{v_i}) = \sqrt{\sum_{j=1}^{i+1} |\langle \vec{v}, \vec{e_j} \rangle|^2} \cdot \vec{e_1} + \sum_{k=i+1}^{n} \langle \vec{v}, \vec{e_k} \rangle \cdot \vec{e_k}$

递推易得 $R_{n-1} \cdots R_1(\vec{v}) = \sqrt{\sum_{j=1}^{n} |\langle \vec{v}, \vec{e_j} \rangle|^2} \cdot \vec{e_1} = \|\vec{v}\| \vec{e_1}$ Q.E.D.

※ 初等旋转在标准正交基下的矩阵：

$$\begin{pmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & \cos\theta & \cdots & -\sin\theta & & \\ & & \vdots & \ddots & \vdots & & \\ & & \sin\theta & \cdots & \cos\theta & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix}$$

- QR分解

$\forall A \in \mathbb{R}^{n,n}, \exists Q, R \in \mathbb{R}^{n,n}$ s.t. $A = QR$.

其中 Q 为正交矩阵，R 为上三角矩阵. (7-78)

我们要证QR分解的加强版：

- $\forall A \in \mathbb{R}^{n,n}, \exists R_1, \cdots, R_N, B \in \mathbb{R}^{n,n}$ s.t. $A = R_1 \cdots R_N B$

其中 $R_1, \cdots, R_N$ 为初等旋转矩阵. B 为上三角矩阵.

且有 $N \leq \frac{1}{2}n(n-1)$. (7-79)

证：对阶数 n 作归纳法

$n=1$ 时显然.

假设上述分解对 $n-1$ 阶矩阵成立；

对 $A \in \mathbb{R}^{n,n}$，用 $n-1$ 个初等旋转把 A 的第一列 $\vec{a_1}$ 变为 $(a, 0, \cdots, 0)^T$.

$\Rightarrow T_{n-1} \cdots T_1 A = \begin{pmatrix} a & * \\ 0 & A_1 \end{pmatrix}$, $A_1 \in \mathbb{R}^{n-1,n-1}$, $T_1, \cdots, T_{n-1}$ 为初等旋转

对 A_1 应用归纳假设：

$A_1 = S_1 \cdots S_M B$；$S_1, \cdots, S_M \in \mathbb{R}^{n-1,n-1}$ 是初等旋转矩阵，

B 是上三角矩阵．$M \le \frac{1}{2}(n-1)(n-2)$．

$$\Rightarrow A = T_1^{-1} \cdots T_{n-1}^{-1} \begin{pmatrix} a_1 & * \\ 0 & S_1 \cdots S_M B \end{pmatrix}$$

$$= T_1^{-1} \cdots T_{n-1}^{-1} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & S_1 \end{pmatrix} \cdots \begin{pmatrix} 1 & 0 \\ 0 & S_M \end{pmatrix}}_{\text{初等旋转}} \underbrace{\begin{pmatrix} a_1 & * \\ 0 & B \end{pmatrix}}_{\text{上三角矩阵}}$$

$$= R_1 \cdots R_N B$$

$N = (n-1) + M \le \frac{1}{2}n(n-1)$． Q.E.D.

• 设 S 是 Euclid 空间 V 中的旋转，$\dim V = n$．
S 可以表示为至多 $\frac{1}{2}n(n-1)$ 个初等旋转的复合． (7-80)

证：在 V 中取标准正交基 $e_1, e_2, \cdots, e_n$，S 在该基下的矩阵为 A．

由 7-79，$\exists$ 至多 $\frac{1}{2}n(n-1)$ 个初等旋转矩阵 $R_1, \cdots, R_N$

s.t. $A = R_1 \cdots R_N B$．

B 是上三角矩阵，下面要证 B 应为单位矩阵 I：

$B = R_N^{-1} \cdots R_1^{-1} A$ 是正交矩阵，因此 B 是正规的．

在复谱定理的证明中，得出了正规的上三角矩阵是对角矩阵．

$\Rightarrow B = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$，$\lambda_i = \pm 1$（等距同构 $|\lambda| = 1$）．

在 7-79 的证明中，总能选取初等旋转使 $\lambda_1, \cdots, \lambda_{n-1} > 0$．

又 $\det B = 1 \Rightarrow \lambda_1 = \cdots = \lambda_n = 1$，$B$ 是单位矩阵．

$\Rightarrow A = R_1 \cdots R_N$． Q.E.D.

※ Euler 角

$\mathbb{R}^3$ 中任意一个刚体定点旋转都可以分解为 3 个初等旋转（定轴转动）．

$\forall S \in SO(3)$，$S(\varphi, \theta, \psi) = R_z(\varphi) R_x(\theta) R_{z'}(\psi)$．

其中：

$$R_z(\varphi)=\begin{pmatrix}\cos\varphi & -\sin\varphi & 0\\ \sin\varphi & \cos\varphi & 0\\ 0 & 0 & 1\end{pmatrix}$$

$$R_x(\theta)=\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\theta & -\sin\theta\\ 0 & \sin\theta & \cos\theta\end{pmatrix}$$

$$R_z'(\psi)=\begin{pmatrix}\cos\psi & -\sin\psi & 0\\ \sin\psi & \cos\psi & 0\\ 0 & 0 & 1\end{pmatrix}$$

(φ,θ,ψ)这样一组用来描述刚体旋转的坐标称为Euler角.

7.8 极分解与奇异值分解

- 正算子

· 自伴算子 $T\in\mathcal{L}(V)$ 称为正算子（半正定算子），若 $\forall\vec{v}\in V$，$\langle T(\vec{v}),\vec{v}\rangle\geqslant 0$.

※ 若V是酉空间，条件（$\forall\vec{v}\in V$，$\langle T(\vec{v}),\vec{v}\rangle\geqslant 0$）已经蕴涵T自伴，见7-47.

※ 若V是Euclid空间，正算子T对应了第6章提到的半正定实二次型.

· 自伴算子T是正算子，当且仅当T的所有本征值非负. (7-81)

证：V上存在标准正交基 $\vec{e}_1,\cdots,\vec{e}_n$，使T成对角形 $\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$（谱定理）.

且 $\lambda_1,\cdots,\lambda_n$ 是T的本征值（$\lambda_1,\cdots,\lambda_n\in\mathbb{R}$）

$\forall\vec{v}\in V$，$T(\vec{v})=T(\sum_{i=1}^n\langle\vec{v},\vec{e}_i\rangle\vec{e}_i)=\sum_{i=1}^n\lambda_i\langle\vec{v},\vec{e}_i\rangle\vec{e}_i$.

$\Rightarrow\langle T(\vec{v}),\vec{v}\rangle=\langle\sum_{i=1}^n\lambda_i\langle\vec{v},\vec{e}_i\rangle\vec{e}_i,\sum_{j=1}^n\langle\vec{v},\vec{e}_j\rangle\vec{e}_j\rangle=\sum_{i=1}^n\lambda_i|\langle\vec{v},\vec{e}_i\rangle|^2$

T是正算子 $\Longleftrightarrow\forall\vec{v}\in V$，$\langle T(\vec{v}),\vec{v}\rangle\geqslant 0$

$\Longleftrightarrow\forall\vec{v}\in V$，$\sum_{i=1}^n\lambda_i|\langle\vec{v},\vec{e}_i\rangle|^2\geqslant 0$

$\Longleftrightarrow\lambda_i\geqslant 0$，$i=1,\cdots,n$. Q.E.D.

· 平方根

算子R称为算子T的平方根，若$R^2=T$.

· 正算子与平方根

设$T\in\mathcal{L}(V)$，以下表述等价： (7-82)

i) T是正算子

ii) T有平方根是正算子

iii) T有平方根是自伴算子

iv) $\exists R\in\mathcal{L}(V)$ s.t. $T=R^*R$

证：i)⇒ii)：设V有由T本征向量构成的标准正交基$\vec{e_1},\cdots,\vec{e_n}$，对应$T$的本征值$\lambda_1,\cdots,\lambda_n$.（$\lambda_i\geqslant 0$）.

设$S\in\mathcal{L}(V)$，$S(\vec{e_i})=\sqrt{\lambda_i}\,\vec{e_i}$.

$\Rightarrow \forall \vec{v}\in V$，$\langle S(\vec{v}),\vec{v}\rangle=\sum_{i=1}^{n}\sqrt{\lambda_i}\,|\langle\vec{v},\vec{e_i}\rangle|^2\geqslant 0$ ⇒ S是正算子.

且$\forall i=1,\cdots,n$，$S^2(\vec{e_i})=\lambda_i\vec{e_i}=T(\vec{e_i})\Rightarrow S^2=T$

ii)⇒iii) 显然.

iii)⇒iv) 设自伴算子R s.t. $T=R^2$，显然$T=R^*R$

iv)⇒i) $\exists R\in\mathcal{L}(V)$ s.t. $T=R^*R$.

$\Rightarrow \forall\vec{v}\in V$，$\langle T(\vec{v}),\vec{v}\rangle=\langle R^*R(\vec{v}),\vec{v}\rangle=\langle R(\vec{v}),R(\vec{v})\rangle\geqslant 0$

⇒ T是正算子. Q.E.D.

※ $\forall T\in\mathcal{L}(V,W)$，$T^*T\in\mathcal{L}(V)$和$TT^*\in\mathcal{L}(W)$都是正算子. (7-83).

· 每个正算子都有唯一的正平方根（平方根是正算子）. (7-84)

证：设R是T的正平方根

V有由R本征向量构成的标准正交基$\vec{e_1},\cdots,\vec{e_n}$，对应$R$的本征值$\varepsilon_1,\cdots,\varepsilon_n$.（$\varepsilon_i\geqslant 0$）.

$\Rightarrow R(\vec{e_i})=\varepsilon_i\vec{e_i}\Rightarrow T(\vec{e_i})=R^2(\vec{e_i})=\varepsilon_i^2\vec{e_i}$.

$\Rightarrow \varepsilon_1^2,\cdots,\varepsilon_n^2$是$T$的本征值. 记$\lambda_i=\varepsilon_i^2$.

$\vec{e_1},\cdots,\vec{e_n}$也是$T$的本征向量，$R=diag(\sqrt{\lambda_1},\cdots,\sqrt{\lambda_n})$显然是唯一的.

※ 若T是正算子，T的唯一正平方根记为$\sqrt{T}$.

- 极分解

· LADR喜欢把$\mathcal{L}(V)$与$\mathbb{C}$作类比.

· 单位圆$\{z\in\mathbb{C}\mid |z|=1\}$是$\mathbb{C}$的重要子集

$|z|=1 \Rightarrow \bar{z}z=1$ 对应$\mathcal{L}(V)$中$T^*T=I$

即$\mathbb{C}$中的单位圆类比$\mathcal{L}(V)$中的等距同构集合.

· $\forall z\in\mathbb{C}$, $z=\frac{z}{|z|}\cdot|z|=\frac{z}{|z|}\cdot\sqrt{\bar{z}z}$ (极坐标表示 $z=\rho e^{i\theta}$)

其中$\frac{z}{|z|}$指示方向，$\sqrt{\bar{z}z}$代表大小.

这个分解类比到$\mathcal{L}(V)$，即算子T等于等距同构与$\sqrt{T^*T}$的复合.

$\sqrt{T^*T}$称为算子T的模量(modulus)，记为$|T|$ (这里不用此记号).

由7-83，T^*T是正算子，因此上述记号是合理的.

· $\forall T\in\mathcal{L}(V)\ (\forall \vec{v}\in V,\ \|T(\vec{v})\|=\|\sqrt{T^*T}(\vec{v})\|)$. (7-85)

证: $\|\sqrt{T^*T}(\vec{v})\|^2=\langle\sqrt{T^*T}(\vec{v}),\sqrt{T^*T}(\vec{v})\rangle=\langle T^*T(\vec{v}),\vec{v}\rangle$ ($\sqrt{T^*T}$自伴)

$=\langle T(\vec{v}),T(\vec{v})\rangle=\|T(\vec{v})\|^2$ Q.E.D.

※ 这说明了算子T与$\sqrt{T^*T}$"大小"相同的特点.

· $\operatorname{Null}T=\operatorname{Null}\sqrt{T^*T}=(\operatorname{Ran}\sqrt{T^*T})^\perp$ (7-86)

证: 第一个等号由7-85立得；第二个等号即7-41. Q.E.D.

· 极分解

设$T\in\mathcal{L}(V)$，存在等距同构$S\in\mathcal{L}(V)$使$T=S\sqrt{T^*T}$. (7-87)

证: 定义S_1: $\operatorname{Ran}\sqrt{T^*T}\to\operatorname{Ran}T$,

$S_1(\sqrt{T^*T}(\vec{v}))=T(\vec{v})$.

S_1是良定义的: 若$\sqrt{T^*T}(\vec{v}_1)=\sqrt{T^*T}(\vec{v}_2)$.

则有$\|T(\vec{v}_1)-T(\vec{v}_2)\|=\|\sqrt{T^*T}(\vec{v}_1)-\sqrt{T^*T}(\vec{v}_2)\|=0$

$\Rightarrow T(\vec{v}_1)=T(\vec{v}_2)$

按以上定义以及7-85，$S_1\in\mathcal{L}(\operatorname{Ran}\sqrt{T^*T},\operatorname{Ran}T)$是等距映射.

$\dim Ran T = \dim Ran\sqrt{T^*T}$，$S_1$是等距同构.

接下来只需要扩充成等距同构 $S \in \mathcal{L}(V)$.

$(Ran\sqrt{T^*T})^\perp = Null T$，$(Ran T)^\perp = Null T^*$.

且有 $\dim Null T = \dim Null T^*$.

则存在等距同构 $S_2 \in \mathcal{L}(Null T, Null T^*)$.

$\forall \vec{v} \in V$，$\vec{v} = \vec{u} + \vec{w}$，其中 $\vec{u} \in Ran\sqrt{T^*T}$，$\vec{w} \in Null T$.

定义 $S: V \to V$，$\forall \vec{v} \in V$，$S(\vec{v}) = S_1(\vec{u}) + S_2(\vec{w})$.

易证 S是线性映射.

$\|S(\vec{v})\|^2 = \|S_1(\vec{u}) + S_2(\vec{w})\|^2 = \|S_1(\vec{u})\|^2 + \|S_2(\vec{w})\|^2 = \|\vec{u}\|^2 + \|\vec{w}\|^2 = \|\vec{v}\|^2$

则 $S \in \mathcal{L}(V)$是等距同构，$T = S\sqrt{T^*T}$ Q.E.D.

※ 酉空间中，等距同构S与正算子$\sqrt{T^*T}$都可以有简单的对角形（一般是相对于两组不同的标准正交基）.

- 奇异值

- 算子T的奇异值是$\sqrt{T^*T}$的本征值（计重数）.

※ 奇异值是非负实数.

$T \in \mathcal{L}(V)$有$\dim V$个奇异值（计重数）.

- 算子T的奇异值是T^*T的本征值的非负平方根（计重数）. (7-88)

证 由谱定理显然.

- T与伴随T^*有相同的奇异值（计重数）. (7-89)

证 $T \in \mathcal{L}(V)$的极分解 $T = S\sqrt{T^*T}$，$S \in \mathcal{L}(V)$是等距同构.

$\Rightarrow TT^* = S\sqrt{T^*T}(S\sqrt{T^*T})^* = S\sqrt{T^*T}\sqrt{T^*T}S^* = S(T^*T)S^{-1}$.

$\exists \vec{v} \neq 0$ s.t. $T^*T(\vec{v}) = \lambda\vec{v} \Rightarrow T^*T(S\vec{v}) = \lambda S(\vec{v})$

$\Rightarrow TT^*(S(\vec{v})) = S(T^*T)S^{-1}(S(\vec{v})) = ST^*T(\vec{v}) = \lambda S(\vec{v})$.

$\Rightarrow TT^*$与T^*T有相同的本征值.

$\Rightarrow T$与T^*有相同的奇异值. Q.E.D.

- Schmidt 分解

· 设$T\in\mathcal{L}(V)$的奇异值为$\sigma_1,\cdots,\sigma_n$. V有两组标准正交基$\vec{v}_1,\cdots,\vec{v}_n$和$\vec{u}_1,\cdots,\vec{u}_n$使$\forall\vec{v}\in V$, $T(\vec{v})=\sum_{i=1}^n\sigma_i<\vec{v},\vec{v}_i>\vec{u}_i$. (7-90)

证 对$\sqrt{T^*T}$应用谱定理：V有标准正交基$\vec{v}_1,\cdots,\vec{v}_n$ s.t. $\sqrt{T^*T}(\vec{v}_i)=\sigma_i\vec{v}_i$

$\forall\vec{v}\in V$, $\vec{v}=\sum_{i=1}^n<\vec{v},\vec{v}_i>\vec{v}_i\Rightarrow\sqrt{T^*T}(\vec{v})=\sum_{i=1}^n\sigma_i<\vec{v},\vec{v}_i>\vec{v}_i$

T有极分解：$T=S\sqrt{T^*T}$.

$\Rightarrow T(\vec{v})=S\sqrt{T^*T}(\vec{v})=S(\sum_{i=1}^n\sigma_i<\vec{v},\vec{v}_i>\vec{v}_i)=\sum_{i=1}^n\sigma_i<\vec{v},\vec{v}_i>S(\vec{v}_i)$.

S是等距同构，则$S(\vec{v}_1),\cdots,S(\vec{v}_n)$是$V$的标准正交基. 令$\vec{u}_i=S(\vec{v}_i)$.

$\Rightarrow\forall\vec{v}\in V$, $T(\vec{v})=\sum_{i=1}^n\sigma_i<\vec{v},\vec{v}_i>\vec{u}_i$ Q.E.D

※ Schmidt分解本质是算子在两组标准正交基下的对角化.

· 设$T\in\mathcal{L}(V)$的奇异值为$\sigma_1,\cdots,\sigma_n$. V有两组标准正交基使T成对角形

$\Sigma=diag(\sigma_1,\cdots,\sigma_n)$ (7-91)

证 T有Schmidt分解 $T(\vec{v})=\sum_{i=1}^n\sigma_i<\vec{v},\vec{v}_i>\vec{u}_i$

$\Rightarrow T(\vec{v}_i)=\sigma_i\vec{u}_i$.

即T关于基$\mathcal{A}=\{\vec{v}_1,\cdots,\vec{v}_n\}$与$\mathcal{B}=\{\vec{u}_1,\cdots,\vec{u}_n\}$的矩阵

$\Sigma=[T]_{\mathcal{A}\mathcal{B}}=diag(\sigma_1,\cdots,\sigma_n)$. Q.E.D.

· 设$T\in\mathcal{L}(V)$有Schmidt分解 $\forall\vec{v}\in V$, $T(\vec{v})=\sum_{i=1}^n\sigma_i<\vec{v},\vec{v}_i>\vec{u}_i$, 则： (7-92)

i) $T^*(\vec{v})=\sum_{i=1}^n\sigma_i<\vec{v},\vec{u}_i>\vec{v}_i$.

ii) 若T可逆，则$T^{-1}(\vec{v})=\sum_{i=1}^n\sigma_i^{-1}<\vec{v},\vec{u}_i>\vec{v}_i$.

证 i) $\forall\vec{v},\vec{w}\in V$, $<\vec{w},T^*(\vec{v})>=<T(\vec{w}),\vec{v}>$

$=<\sum_{i=1}^n\sigma_i<\vec{w},\vec{v}_i>\vec{u}_i,\vec{v}>=\sum_{i=1}^n\sigma_i<\vec{w},\vec{v}_i><\vec{u}_i,\vec{v}>$

$=\sum_{i=1}^n\overline{\sigma_i<\vec{v},\vec{u}_i>}<\vec{w},\vec{v}_i>=<\vec{w},\sum_{i=1}^n\sigma_i<\vec{v},\vec{u}_i>\vec{v}_i>$

$\Rightarrow T^*(\vec{v})=\sum_{i=1}^n\sigma_i<\vec{v},\vec{u}_i>\vec{v}_i$

ii) 若T可逆，易证$\sqrt{T^*T}$可逆，$\sqrt{T^*T}$本征值非零(否则$\mathrm{Null}\sqrt{T^*T}\neq\{0\}$)

即$\sigma_1,\cdots,\sigma_n\neq0$.

$$令\ \vec{w} = \sum_{i=1}^{n} \sigma_i^{-1} \langle \vec{v}, \vec{u}_i \rangle \vec{v}_i$$

$$\Rightarrow T(\vec{w}) = \sum_{i=1}^{n} \sigma_i^{-1} \langle \vec{v}, \vec{u}_i \rangle T(\vec{v}_i) = \sum_{i=1}^{n} \sigma_i^{-1} \langle \vec{v}, \vec{u}_i \rangle \sigma_i \vec{u}_i$$

$$= \sum_{i=1}^{n} \langle \vec{v}, \vec{u}_i \rangle \vec{u}_i = \vec{v}$$

$$\Rightarrow T^{-1}(\vec{v}) = \vec{w} = \sum_{i=1}^{n} \sigma_i^{-1} \langle \vec{v}, \vec{u}_i \rangle \vec{v}_i$$ Q.E.D.

- 奇异值分解(SVD)

- (Compact Form)

设 $A \in \mathbb{F}^{n,n}$. 存在酉矩阵 $U, V \in \mathbb{F}^{n,n}$ 使 $A = U\Sigma V^*$,

其中 $\Sigma = \mathrm{diag}(\sigma_1, \cdots, \sigma_n)$ 是由 A 的奇异值构成的对角矩阵. (7-93)

证: $\mathbb{F}^{n,n}$ 上有标准正交基 $\vec{v}_1, \cdots, \vec{v}_n$ 与 $\vec{u}_1, \cdots, \vec{u}_n$ 使 $A\vec{v}_i = \sigma_i \vec{u}_i$

令 $V = (\vec{v}_1 \cdots \vec{v}_n)$, $U = (\vec{u}_1 \cdots \vec{u}_n)$ 是酉矩阵.

$$\Rightarrow AV = A(\vec{v}_1 \cdots \vec{v}_n) = (A\vec{v}_1 \cdots A\vec{v}_n) = (\sigma_1 \vec{u}_1 \cdots \sigma_n \vec{u}_n) = U\Sigma$$

$$\Rightarrow A = U\Sigma V^{-1} = U\Sigma V^*$$ Q.E.D.

- (Original Form)

设 $A \in \mathbb{F}^{m,n}$, 即使 $m \neq n$, 对 A 仍可定义奇异值. 设 A 的非零奇异值为 $\sigma_1, \cdots, \sigma_r$ $(r = \mathrm{rank} A)$, 存在酉矩阵 $U \in \mathbb{F}^{m,n}$ 和 $V \in \mathbb{F}^{n,n}$ 使 $A = U\Sigma V^*$. 其中 $\Sigma \in \mathbb{F}^{m,n}$ 是如下"对角矩阵":

$$\Sigma_{ij} = \begin{cases} \sigma_i, & i = j \leq r \\ 0, & i \neq j \end{cases} \tag{7-94}$$

证: 对 A 作 Schmidt 分解. 存在 $\mathbb{F}^n$ 上标准正交基 $\vec{v}_1, \cdots, \vec{v}_n$ 与 $\mathbb{F}^m$ 上标准正交基 $\vec{u}_1, \cdots, \vec{u}_m$, 使 $A\vec{v}_i = \begin{cases} \sigma_i \vec{u}_i, & i \leq r. \\ 0, & i > r. \end{cases}$

$$\Rightarrow AV = (A\vec{v}_1 \cdots A\vec{v}_n) = (\sigma_1 \vec{u}_1 \cdots \sigma \vec{u}_r\ 0 \cdots 0) = U\Sigma$$

$$\Rightarrow A = U\Sigma V^*$$ Q.E.D.

※ 奇异值分解到极分解:

$$\underset{\substack{\uparrow \\ T的矩阵}}{A} = U\Sigma V^* = \underbrace{(UV^*)}_{S的矩阵}\underbrace{(V\Sigma V^*)}$$

▸ 算子范数

· 设$T\in\mathcal{L}(V)$，$\hat{\sigma}$为T的最小奇异值，σ为T的最大奇异值，则： (7-95)

i) $\forall\vec{v}\in V$，$\hat{\sigma}\|\vec{v}\|\leq\|T(\vec{v})\|\leq\sigma\|\vec{v}\|$.

ii) 若λ是T的本征值，有$\hat{\sigma}\leq|\lambda|\leq\sigma$.

证 i) Schmidt分解：$T(\vec{v})=\sum_{i=1}^{n}\sigma_i\langle\vec{v},\vec{v}_i\rangle\vec{u}_i$

$\Rightarrow\|T(\vec{v})\|^2=\sum_{i=1}^{n}\sigma_i^2|\langle\vec{v},\vec{v}_i\rangle|^2$

$(\hat{\sigma}\|\vec{v}\|)^2=\hat{\sigma}^2\sum_{i=1}^{n}|\langle\vec{v},\vec{v}_i\rangle|^2\leq\sum_{i=1}^{n}\sigma_i^2|\langle\vec{v},\vec{v}_i\rangle|^2$

$\leq\sigma^2\sum_{i=1}^{n}|\langle\vec{v},\vec{v}_i\rangle|^2=(\sigma\|\vec{v}\|)^2$

$\Rightarrow\hat{\sigma}\|\vec{v}\|\leq\|T(\vec{v})\|\leq\sigma\|\vec{v}\|$.

设T的Schmidt分解：$T(\vec{e}_i)=\sigma_i\vec{f}_i$

则左边等号成立当且仅当$\vec{v}=\hat{\vec{e}}$，右边等号成立当且仅当$\vec{v}=\vec{e}$.

ii) 设$\vec{v}$是λ对应的本征向量. $\|T(\vec{v})\|=|\lambda|\cdot\|\vec{v}\|$.

$\Rightarrow\hat{\sigma}\|\vec{v}\|\leq|\lambda|\cdot\|\vec{v}\|\leq\sigma\|\vec{v}\|\Rightarrow\hat{\sigma}\leq|\lambda|\leq\sigma$ Q.E.D.

· V上的闭单位球$S=\{\vec{v}\in V\mid\|\vec{v}\|\leq1\}$.

· 定义映射$\|\cdot\|_{op}:\mathcal{L}(V)\to\mathbb{R}$，$\forall T\in\mathcal{L}(V)$，$\|T\|_{op}=\max\{\|T(\vec{v})\|\mid\vec{v}\in S\}$

称之为算子范数，它使$\mathcal{L}(V)$成为赋范向量空间.

· 设σ为T的最大奇异值，则$\|T\|_{op}=\sigma$. (7-96)

证 由(7-95)显然

· $\|\cdot\|_{op}$满足范数的性质： (7-97)

i) 正定性：$\forall T\in\mathcal{L}(V)$，$\|T\|_{op}\geq0$；$\|T\|_{op}=0\iff T=0$.

ii) 齐性：$\forall\lambda\in\mathbb{F}$，$T\in\mathcal{L}(V)$，$\|\lambda T\|_{op}=|\lambda|\cdot\|T\|_{op}$.

iii) 三角不等式：$\forall S,T\in\mathcal{L}(V)$，$\|S+T\|_{op}\leq\|S\|_{op}+\|T\|_{op}$

证 设T的奇异值为$\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_n\geq0$.

i) $\|T\|_{op}=\sigma_1\geq0$.

$\|T\|_{op}=0\iff\sigma_1=\cdots=\sigma_n=0\iff\sqrt{T^*T}=0\iff T=0$.

ii) T^*T 的本征值为 $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_n^2 \geq 0$.

$\Rightarrow \bar{\lambda}T^* \cdot \lambda T$ 的本征值为 $|\lambda|^2\sigma_1^2 \geq |\lambda|^2\sigma_2^2 \geq \cdots \geq |\lambda|^2\sigma_n^2 \geq 0$.

$\Rightarrow \lambda T$ 的奇异值为 $|\lambda|\sigma_1 \geq |\lambda|\sigma_2 \geq \cdots |\lambda|\sigma_n \geq 0$.

$\Rightarrow \|\lambda T\|_{op} = |\lambda|\sigma_1 = |\lambda| \cdot \|T\|_{op}$

iii) 设 $U = S + T$ 的最大奇异值为 σ_u.

S, T 的最大奇异值为 σ_s, σ_t.

$\exists \vec{v} \in V$ s.t. $\|\vec{v}\| = 1$ & $\sqrt{U^*U}(\vec{v}) = \sigma_u \vec{v}$

$\Rightarrow \sigma_u = \|\sigma_u \vec{v}\| = \|\sqrt{U^*U}(\vec{v})\| = \|U(\vec{v})\| = \|S(\vec{v}) + T(\vec{v})\|$

$\leq \|S(\vec{v})\| + \|T(\vec{v})\| \leq \sigma_s \|\vec{v}\| + \sigma_t \|\vec{v}\| = \sigma_s + \sigma_t$.

于是 $\|S+T\|_{op} \leq \|S\|_{op} + \|T\|_{op}$. Q.E.D.

※ 算子还可以定义其它的范数.

比如 $\langle S, T \rangle = tr(S^*T)$ 是 $\mathcal{L}(V)$ 上的内积，它诱导的范数 $\|T\|_2 = \sqrt{tr(T^*T)} = \sqrt{\sum_{i=1}^{n} \sigma_i^2}$ 称为 Frobenius 范数或 Hilbert-Schmidt 范数.

由 $\|T\|_{op} = \sigma_{max}$, $\|T\|_2 = \sqrt{\sum_{i=1}^{n} \sigma_i^2}$ 有：

$\forall T \in \mathcal{L}(V)$, $\|T\|_{op} \leq \|T\|_2$

重线性代数

8.1 对偶空间

▸ 向量空间的对偶

· 线性泛函：从向量空间到数域的映射.

· 对偶空间

V上全体线性泛函构成的集合$\mathcal{L}(V,\mathbb{F})$称为$V$的对偶空间，

记为$V'=\mathcal{L}(V,\mathbb{F})$

· 有限维向量空间与对偶维数相同（向量空间与对偶同构） (8-1)

证 $\dim V'=\dim\mathcal{L}(V,\mathbb{F})=\dim V\cdot\dim\mathbb{F}=\dim V$ □

· 对偶基

设$v_1,\cdots,v_n$是V的基，$\exists\varphi_1,\cdots,\varphi_n\in V'$ s.t. $\varphi_i(v_j)=\delta_{ij}$.

$\varphi_1,\cdots,\varphi_n$是$V'$的基，称为$v_1,\cdots,v_n$的对偶基. (8-2)

证 $\exists\lambda_1,\cdots,\lambda_n\in\mathbb{F}$ s.t. $\sum_{i=1}^n\lambda_i\varphi_i=0$.

$\forall j=1,\cdots,n$. $(\sum_{i=1}^n\lambda_i\varphi_i)(v_j)=\sum_{i=1}^n\lambda_i\delta_{ij}=\lambda_j\Rightarrow\lambda_j=0$.

$\Rightarrow\lambda_1=\cdots=\lambda_n=0$. $\varphi_1,\cdots,\varphi_n$是$V'$中的线性无关组

又$\dim V'=n\Rightarrow\varphi_1,\cdots,\varphi_n$是$V'$的基. □

※ $\forall v\in V$, $v=\sum_{i=1}^n\lambda_iv_i\Rightarrow\varphi_i(v)=\lambda_i\Rightarrow v=\sum_{i=1}^n\varphi_i(v)v_i$ (8-3)

$\forall\varphi\in V'$, $\varphi=\sum_{i=1}^n\varepsilon_i\varphi_i\Rightarrow\varphi(v_i)=\varepsilon_i\Rightarrow\varphi=\sum_{i=1}^n\varphi(v_i)\varphi_i$ (8-4)

于是向量空间的坐标映射：

$v\mapsto(x_1,\cdots,x_n)^T=(\varphi_1,\cdots,\varphi_n)^T(v)$ (8-5)

本质上就是把对偶基作用在向量上.

· 对偶空间的过渡矩阵

设$v_1,\cdots,v_n$与$u_1,\cdots,u_n$是V的两组基，对偶基为$\varphi_1,\cdots,\varphi_n$与$\psi_1,\cdots,\psi_n$.

$\forall v\in V$, $v=\sum_{i=1}^n\varphi_i(v)v_i=\sum_{i=1}^n\psi_i(v)u_i$

$\Rightarrow(v_1\cdots v_n)(\varphi_1\cdots\varphi_n)^T=(u_1\cdots u_n)(\psi_1\cdots\psi_n)^T$.

若P是$v_1,\cdots,v_n$到$u_1,\cdots,u_n$的过渡矩阵，易得$(P^{-1})^T$是$\varphi_1,\cdots,\varphi_n$到

$\varphi_1,\cdots,\varphi_n$的过渡矩阵. (8-6)

- 双重对偶

设V'是V的对偶空间,V''是V'的对偶空间. 考虑映射$\Lambda: V\to V''$,
$v\mapsto \Lambda_v$ s.t. $\forall \varphi\in V'$, $\Lambda_v(\varphi)=\varphi(v)$. (8-7)

i) $\Lambda\in\mathcal{L}(V,V'')$

ii) 若V是有限维向量空间,则Λ是一个同构.

证 i) $\forall v,u\in V$, $a,b\in\mathbb{F}$:

$\forall \varphi\in V'$, $\Lambda(au+bv)(\varphi)=\varphi(au+bv)=a\varphi(u)+b\varphi(v)$

$=a\Lambda(u)(\varphi)+b\Lambda(v)(\varphi)$

$=(a\Lambda(u)+b\Lambda(v))(\varphi)$

$\Rightarrow \Lambda(au+bv)=a\Lambda(u)+b\Lambda(v)$, Λ是线性映射.

ii) $\forall v\in \mathrm{Null}\,\Lambda$, $\Lambda_v=\Lambda(v)=0$.

$\Rightarrow \forall\varphi\in V'$, $\varphi(v)=0(\varphi)=0$.

利用8-3展开: $v=\sum_{i=1}^n\varphi_i(v)v_i=\sum_{i=1}^n 0\cdot v_i=0\Rightarrow \mathrm{Null}\,\Lambda=\{0\}$.

又$\dim V''=\dim V'=\dim V$, $\Lambda\in\mathcal{L}(V,V'')$是一个同构. □

※ Λ是有限维V到V''的同构,它不依赖于基的选取,因此是一个"自然同构"(canonical isophorism). 向量空间V与V''是自然同构的,且$V\cap V''=\varnothing$

因此可以把V与V''等同起来,视为一个向量空间.

即: $(V')'=V$; $\forall v\in V, v'\in V'$, $v'(v)=v(v')$. (8-8)

V与V'互为对偶空间. (仅限有限维!)

- 内积空间的对偶

- Riesz表示定理(重新提出).

设V是有限维内积空间,$\forall\varphi\in V'$, $\exists$唯一$u\in V$ s.t. $\forall v\in V$,
$\varphi(v)=\langle v,u\rangle$. (7-26)

i.e. $u\mapsto\varphi$是V与V'之间的一个双射.

· $f: V \to V'$，$f(u)=\varphi$ 由前述定义，f 是共轭线性的． (8-9)

证 $\forall u_1, u_2 \in V$，$f(u_1)=\varphi_1$，$f(u_2)=\varphi_2$．

$\forall v \in V$，$\varphi_1(v) = \langle v, u_1\rangle$，$\varphi_2(v) = \langle v, u_2 \rangle$．

$\Rightarrow (\varphi_1+\varphi_2)(v) = \langle v, u_1+u_2\rangle \Rightarrow f(u_1)+f(u_2)=f(u_1+u_2)$．

$\forall \lambda \in F$，$u \in V$，$f(u)=\varphi$．

$\forall v \in V$，$\varphi(v) = \langle v, u\rangle \Rightarrow \langle v, \lambda u\rangle = \bar{\lambda}\varphi(v)$

$\Rightarrow f(\lambda u) = \bar{\lambda} f(u)$． □

※ 若 V 是 Euclid 空间，则 Riesz 表示定理给出了一个 V 与 V' 的自然同构（$\mathbb{R}$ 上共轭线性就是线性）；若 V 是酉空间，Riesz 表示定理只给出了 V 与 V' 的自然双射，但不是同构．

· 线性映射的对偶

· 对偶映射

设 $T \in \mathcal{L}(V, W)$，T 的对偶映射 $T' \in \mathcal{L}(W', V')$ 定义为：

$\forall \varphi \in W'$，$T'(\varphi) = \varphi \circ T$

· 对偶映射对应矩阵转置．

设 V 有基 $v_1, \cdots, v_n$ 及 V' 有对偶基 $\varphi_1, \cdots, \varphi_n$；$W$ 有基 $w_1, \cdots, w_m$ 及 W' 有对偶基 $\psi_1, \cdots, \psi_m$．若 $T \in \mathcal{L}(V, W)$ 在相应基下矩阵为 A，则 $T' \in \mathcal{L}(W', V')$ 在相应基下的矩阵为 A^T． (8-10)

证 设 T 矩阵为 $A=\{a_{i,j}\}$，T' 矩阵为 $B=\{b_{i,j}\}$

$T'(\psi_j) = \sum_{r=1}^{n} b_{r,j}\varphi_r = \psi_j \circ T$

$\Rightarrow \psi_j \circ T(v_i) = \sum_{r=1}^{n} b_{r,j}\varphi_r(v_i) = \sum_{r=1}^{n} b_{r,j}\delta_{r,i} = b_{i,j}$．

又 $T(v_i) = \sum_{r=1}^{m} a_{r,i} w_r$

$\Rightarrow \psi_j \circ T(v_i) = \sum_{r=1}^{m} a_{r,i}\psi_j(w_r) = \sum_{r=1}^{m} a_{r,i}\delta_{j,r} = a_{j,i}$

$\Rightarrow a_{j,i} = b_{i,j} \Rightarrow B = A^T$ □

· 对偶映射的性质． (8-11)

i) $\forall S,T\in\mathcal{L}(V,W)$, $(S+T)'=S'+T'$

ii) $\forall\lambda\in\mathbb{F}, T\in\mathcal{L}(V,W)$, $(\lambda T)'=\lambda T'$

iii) $\forall S\in\mathcal{L}(V,W), T\in\mathcal{L}(U,V)$, $(ST)'=T'S'$

iv) $\forall T\in\mathcal{L}(V,W)$, $(T')'=T$

- §4.6与§7.4给出了矩阵基本子空间的若干结论.
  接下来从对偶映射与零化子的角度重新理解.
- 零化子
  设$U\subset V$, U的零化子定义为$U^0=\{\varphi\in V'\mid\forall u\in U, \varphi(u)=0\}$
- 零化子U^0是V'的子空间. (8-12)

证 $\forall\varphi,\psi\in U^0$, 有$\forall u\in U$, $\varphi(u)=0$, $\psi(u)=0$.

$\Rightarrow\forall u\in U$, $(\varphi+\psi)(u)=0\Rightarrow\varphi+\psi\in U^0$

$\forall\varphi\in U^0$, $\lambda\in\mathbb{F}$, 有$\forall u\in U$, $\varphi(u)=0\Rightarrow\lambda\varphi(u)=0\Rightarrow\lambda\varphi\in U^0$

$\forall u\in U$, $0(u)=0\Rightarrow 0\in U^0$. □

- 若U是有限维向量空间V的子空间, 则$\dim V=\dim U+\dim U^0$ (8-13)

证 设$i\in\mathcal{L}(U,V)$ s.t. $\forall u\in U$, $i(u)=u$

$\Rightarrow i'\in\mathcal{L}(V',U')$, $\operatorname{Null} i'=U^0$

由线性映射基本定理: $\dim V'=\dim\operatorname{Null} i'+\dim\operatorname{Ran} i'$

$\Rightarrow\dim V=\dim U^0+\dim\operatorname{Ran} i'$

下面证$\operatorname{Ran} i'=U'$:

$\forall\varphi\in U'$, $\exists\psi\in V'$ s.t. $\forall u\in U$, $\psi(u)=\varphi(u)$.

$\Rightarrow\psi\circ i(u)=\varphi(u)\Rightarrow i'(\psi)=\varphi\in\operatorname{Ran} i'$

$\Rightarrow U'\subset\operatorname{Ran} i'$. 又$\operatorname{Ran} i'\subset U'\Rightarrow\operatorname{Ran} i'=U'$

$\Rightarrow\dim V=\dim U^0+\dim U'=\dim U+\dim U^0$ □

- T'的零空间:

$\operatorname{Null} T'=(\operatorname{Ran} T)^0$ (8-14)

证 i) $\forall \varphi \in \mathrm{Null}\, T'$, $\varphi \circ T = T'(\varphi) = 0$.

$\Rightarrow \forall v \in V$, $\varphi(T(v)) = \varphi \circ T(v) = 0$

$\Rightarrow \varphi \in (\mathrm{Ran}\, T)^0 \Rightarrow \mathrm{Null}\, T' \subset (\mathrm{Ran}\, T)^0$

ii) $\forall \varphi \in (\mathrm{Ran}\, T)^0$, 有 $\forall v \in V$, $\varphi(T(v)) = 0$

$\Rightarrow T'(\varphi) = \varphi \circ T = 0 \Rightarrow \varphi \in \mathrm{Null}\, T' \Rightarrow (\mathrm{Ran}\, T)^0 \subset \mathrm{Null}\, T'$

$\Rightarrow \mathrm{Null}\, T' = (\mathrm{Ran}\, T)^0$ □

- 设 $T \in \mathcal{L}(V, W)$, V 与 W 有限维, 则

$$\dim \mathrm{Null}\, T' = \dim \mathrm{Null}\, T + \dim W - \dim V \quad (8\text{-}15)$$

证 $\dim \mathrm{Null}\, T' = \dim (\mathrm{Ran}\, T)^0$

$= \dim W - \dim \mathrm{Ran}\, T$

$= \dim W - (\dim V - \dim \mathrm{Null}\, T)$ □

- T' 的值域

设 $T \in \mathcal{L}(V, W)$, V 与 W 有限维, 则: (8-16)

i) $\dim \mathrm{Ran}\, T' = \dim \mathrm{Ran}\, T$

ii) $\mathrm{Ran}\, T' = (\mathrm{Null}\, T)^0$

证 i) 由线性映射基本定理:

$\dim \mathrm{Ran}\, T' = \dim W' - \dim \mathrm{Null}\, T'$

$= \dim W - \dim (\mathrm{Ran}\, T)^0$

$= \dim \mathrm{Ran}\, T$

ii) $\forall \varphi \in \mathrm{Ran}\, T'$, $\exists \psi \in W'$ s.t. $\varphi = T'(\psi)$.

$\forall v \in \mathrm{Null}$, $\varphi(v) = (T'(\psi))(v) = \psi \circ T(v) = \psi(0) = 0$.

$\Rightarrow \varphi \in (\mathrm{Null}\, T)^0 \Rightarrow \mathrm{Ran}\, T' \subset (\mathrm{Null}\, T)^0$

$\dim \mathrm{Ran}\, T' = \dim \mathrm{Ran}\, T$

$= \dim V - \dim \mathrm{Null}\, T$

$= \dim (\mathrm{Null}\, T)^0$

$\Rightarrow Ran\,T' = (Null\,T)^0$. □

- T是单(满)射当且仅当T'是满(单)射. (8-17)

证 T是单射$\Leftrightarrow Null\,T = \{0\} \Leftrightarrow \dim Ran\,T = \dim V$

$\Leftrightarrow \dim Ran\,T' = \dim V' \Leftrightarrow Ran\,T' = V'$

$\Leftrightarrow T'$是满射.

另一方面, $(T')' = T$, T'是单射$\Leftrightarrow T$是满射. □

- 矩阵的基本子空间 again. (8-18)

$\forall A \in \mathbb{F}^{m,n}$, $Null\,A$, $Null\,A^T$, $Ran\,A$, $Ran\,A^T$有如下关系:

i) $Null\,A^T = (Ran\,A)^0$

ii) $Ran\,A^T = (Null\,A)^0$

iii) $\dim Null^T = \dim Null\,T + m - n$.

iv) $\dim Ran\,A = \dim Ran\,A^T = rank\,A$ (行秩等于列秩!)

▸ Euclid空间的一些注解.

- 由Riesz表示定理, Euclid空间V上的内积诱导了V与V'之间的一个自然同构, 那么可以把V与V'等同起来 (记为$V \cong V'$).
- $\forall \varphi \in V'$, $\exists u \in V$ s.t. $\forall v \in V$, $\varphi(v) = \langle v, u \rangle$.

那么把φ视为V中的元素u, 即 $\varphi(v) := \langle v, \varphi \rangle$, $\varphi \in V$. (8-19)

- Euclid空间V的一组基$v_1, \cdots, v_n$的对偶基是V的另一组基$\varphi_1, \cdots, \varphi_n$ s.t. $\langle v_i, \varphi_j \rangle = \delta_{i,j}$.
- Euclid空间V的一组基是标准正交基, 当且仅当它的对偶基等于自身
- Euclid空间之间的映射的对偶$T' \in \mathcal{L}(W', V')$, 也即映射的伴随$T^* \in \mathcal{L}(W, V)$. (8-20)
- Euclid空间V的子空间U的零化子U^0就是U的正交补$U^\perp$. (8-21)

※ 这种把自然同构视为等同的作法在之后有更多应用.

8.2 张量积

- ▸ 多重线性映射

- · 设 $V_1,\cdots,V_p$；W 是向量空间，映射 $f: V_1\times\cdots\times V_p\to W$ 称为多线性映射，若 $\forall\lambda,\mu\in\mathbb{F}$ 以及 $\forall i=1,\cdots,p$ 有：

$$f(v_1,\cdots,\lambda u_i+\mu v_i,\cdots,v_r)=\lambda f(v_1,\cdots,u_i,\cdots,v_r)+\mu f(v_1,\cdots,v_i,\cdots,v_r).$$

多线性映射的全体记为 $Hom(V_1,\cdots,V_p;W)$ 或 $\mathcal{L}(V_1,\cdots,V_p;W)$，后者其实是对记号 $\mathcal{L}$ 的滥用.

※ 双线性形式是 $V\times V\to\mathbb{F}$ 的多线性映射.

- · 关于直积 $V_1\times\cdots\times V_p$.

在 §4.9 积空间中，我们给 $V_1\times\cdots\times V_p$ 赋予加法与数乘，使之成向量空间. 有 $\dim(V_1\times\cdots\times V_p)=\sum_{i=1}^{p}\dim V_i$

在多重线性映射中，$V_1\times\cdots\times V_p$ 是由直积构成的形式集合，其上的元素之间没有运算（事实上 $\mathcal{L}(V_1,\cdots,V_p;W)$ 与线性映射 $\mathcal{L}(V_1\otimes\cdots\otimes V_p,W)$ 有自然同构，见后面的讨论）

简单地理解，直积 ≠ 积空间；直积 ≅ 张量积空间.

- ▸ 张量积作为形式积

- · 张量积是最广义的多线性映射

- · <定义 I> $\forall v_i\in V_i$，张量积 $v_1\otimes\cdots\otimes v_p$ 是满足多线性的"纯粹形式积"：

$\forall\lambda,\mu\in\mathbb{F}$，$i=1,\cdots,p$，

$$v_1\otimes\cdots\otimes(\lambda u_i+\mu v_i)\otimes\cdots\otimes v_p=\lambda(v_1\otimes\cdots\otimes u_i\otimes\cdots\otimes v_p)+\mu(v_1\otimes\cdots\otimes v_i\otimes\cdots\otimes v_p)$$

- · 全体 $v_1\otimes\cdots\otimes v_p$ 张成的向量空间记为 $V_1\otimes\cdots\otimes V_p$.

$V_1\otimes\cdots\otimes V_p$ 称为 $V_1,\cdots,V_p$ 的张量积，其元素称为张量.

- · 向量的张量积也相当于如下多线性映射：

$$\tau: V_1\times\cdots\times V_p\longrightarrow V_1\otimes\cdots\otimes V_p,\ (v_1,\cdots,v_p)\longmapsto v_1\otimes\cdots\otimes v_p \qquad (8\text{-}23)$$

※ 作为最广义的多线性映射，张量积上附加结构成为其它多线性映射，这

表明张量积的万有性．下面给出正式定义：

- 张量积的万有性．

· 为便于讨论，先考虑两个向量空间的情况．

<定义Ⅱ> 设 U, V 是向量空间，向量空间 T 配上双线性映射 $\tau \in \mathcal{L}(U, V; T)$ 称为 U, V 上的张量积 (T, τ)，若有：

$\forall$ 向量空间 W & 双线性映射 $f \in \mathcal{L}(U, V; W)$，

$\exists$ 唯一 $\varphi \in \mathcal{L}(T, W)$ s.t. $f = \varphi \circ \tau$．

$$\begin{tikzcd} U \times V \arrow[r, "\tau"] \arrow[d, "f"'] & T \arrow[dl, dashed, "\varphi"] \\ W & \end{tikzcd}$$

· 张量积在同构的意义下是唯一的．

若 (T, τ) 是 U, V 的张量积，$(\tilde{T}, \tilde{\tau})$ 也是 U, V 的张量积当且仅当 $\exists$ 唯一的同构 $\varphi \in \mathcal{L}(T, \tilde{T})$ s.t. $\tilde{\tau} = \varphi \circ \tau$ (8-24)

$$\begin{tikzcd} U \times V \arrow[r, "\tau"] \arrow[d, "\tilde{\tau}"'] & T \arrow[dl, dashed, shift left, "\varphi"] \\ \tilde{T} \arrow[ur, dashed, shift left, "\tilde{\varphi}"] & \end{tikzcd}$$

证 必要性：若 (T, τ) 是张量积，

$\exists$ 唯一 $\varphi \in \mathcal{L}(T, \tilde{T})$ s.t. $\tilde{\tau} = \varphi \circ \tau$（见右图）

下面只要证 φ 是同构（可逆映射）．

$(\tilde{T}, \tilde{\tau})$ 是张量积，则 $\exists$ 唯一 $\tilde{\varphi} \in \mathcal{L}(\tilde{T}, T)$ s.t. $\tau = \tilde{\varphi} \circ \tilde{\tau}$

$\tau = \tilde{\varphi} \circ \tilde{\tau} = \tilde{\varphi} \circ \varphi \circ \tau \Rightarrow \tilde{\varphi} \circ \varphi = I_T$

$\tilde{\tau} = \varphi \circ \tau = \varphi \circ \tilde{\varphi} \circ \tilde{\tau} \Rightarrow \varphi \circ \tilde{\varphi} = I_{\tilde{T}}$

$\Rightarrow \tilde{\varphi} = \varphi^{-1}$．$\varphi$ 是同构．

充分性：设有向量空间 W 以及双线性映射 $f: U \times V \to W$．

若 (T, τ) 是张量积，则 $\exists$ 唯一 $\rho \in \mathcal{L}(T, W)$．

s.t. $f = \rho \circ \tau$．

$$\begin{tikzcd} & \tilde{T} & \\ U \times V \arrow[ur, "\tilde{\tau}"] \arrow[r, "\tau"] \arrow[d, "f"'] & T \arrow[u, dashed, "\varphi"'] \arrow[dl, dashed, "\rho"] \\ W & \end{tikzcd}$$

若 $\varphi \in \mathcal{L}(T, \tilde{T})$ 是唯一同构，则 $\exists$ 唯一 $\tilde{\rho} = \rho \circ \varphi^{-1}$

s.t. $f = \rho \circ \varphi^{-1} \circ \tilde{\tau} = \tilde{\rho} \circ \tilde{\tau}$．即 $(\tilde{T}, \tilde{\tau})$ 也是张量积． □

※ U, V 的张量积 T 一般记为 $U \otimes V$．

像 $\tau(u, v)$ 记为 $u \otimes v$，$\tau \in \mathcal{L}(U, V; U \otimes V)$ (8-25)

与定义Ⅰ的 8-23 式一致．

· 张量积的唯一性已证. 然而任两个向量空间是否有张量积尚不明确. 下面给出张量积存在性的一个构造性证明. 构造的张量积可作为一个新的定义.

· <定义Ⅲ> 设 U, V 是向量空间, 直积 $U\times V=\{(u,v)\mid u\in U, v\in V\}$.

以 $U\times V$ 的全体元素为基张成的无限维向量空间为 M.

M 的元素是有限个 (u,v) 的线性组合 $\sum_i \lambda_i(u_i,v_i)$.

设 M 的子空间 N 由全体如下形式的向量张成:

$(\lambda u_1+\mu u_2, v)-\lambda(u_1,v)-\mu(u_2,v)$;

$(u, \lambda v_1+\mu v_2)-\lambda(u,v_1)-\mu(u,v_2)$. $\quad\forall\lambda,\mu\in\mathbb{F}$.

商空间 $T=M/N$, 配上映射 $\tau: U\times V\to T$, $(u,v)\mapsto(u,v)+N$.

称为 U, V 的(典型)张量积 (T,τ).

T 记为 $U\otimes V$; $\tau(u,v)=(u,v)+N$ 记为 $u\otimes v$.

· 定义Ⅲ构造的(典型)张量积满足定义Ⅱ(万有性). (8-26)

证 各集合映射关系如右图.

$$\begin{tikzcd} U\times V \arrow[rr, "\tau"] \arrow[dr, "j"] \arrow[ddr, "f"'] & & M/N \arrow[ddl, dashed, "\tilde\sigma"] \\ & M \arrow[ur, "\pi"'] \arrow[d, dashed, "\sigma"] & \\ & W & \end{tikzcd}$$

易证 $\forall f\in\mathcal{L}(U,V;W)$, $\exists$ 唯一的 $\sigma\in\mathcal{L}(M,W)$ s.t. $f=\sigma\circ j$. (但 j 不是双线性的).

映射 σ 把 N 的所有张成元都映到 0:

$\sigma[(\lambda u_1+\mu u_2,v)-\lambda(u_1,v)-\mu(u_2,v)]$

$\quad= f(\lambda u_1+\mu u_2,v)-\lambda f(u_1,v)-\mu f(u_2,v)] = 0$, 另一个同理.

于是 $N\subset \operatorname{Null}\sigma$.

$\forall (u,v)+N\in M/N$, 设 $\tilde\sigma((u,v)+N)=\sigma(\overset{M中的向量}{(u,v)}) = f(\overset{U\times V中的元素}{(u,v)})$

(这里的 $\tilde\sigma$ 是良定义的, 见§4.9 P67).

再证 $\tilde\sigma$ 的唯一性:

设有 $\tilde\sigma'\in\mathcal{L}(M/N,W)$ 使得 $f=\tilde\sigma'\circ\tau$; $\sigma'=\tilde\sigma'\circ\pi$, $f=\sigma'\circ j$.

其中唯一性使 $\sigma=\sigma'$. $\Rightarrow \tilde\sigma'\circ\pi=\tilde\sigma\circ\pi \Rightarrow \tilde\sigma'=\tilde\sigma$.

即有：$\forall f\in \mathcal{L}(U,V;W)$，$\exists$ 唯一 $\tilde{\sigma}\in \mathcal{L}(M/N,W)$ s.t. $f=\tilde{\sigma}\circ\tau$

$(M/N,\tau)$ 是满足定义Ⅱ的张量积． □

※ 以上给出了张量积存在性的不依赖于基的构造．而对有限维的情形，考虑基在张量积中的像，证明是 trivial 的．

- $U\times V$ 上的双线性对应 $U\otimes V$ 上的线性．

即有自然同构 $\mathcal{L}(U,V;W)\cong \mathcal{L}(U\otimes V,W)$，$\sigma(f)(u\otimes v)=f(u,v)$． (8-27)

证 ① σ 是线性的：

$\forall f,g\in\mathcal{L}(U,V;W)$；$\lambda,\mu\in\mathbb{F}$，

$$\begin{aligned}\sigma(\lambda f+\mu g)(u\otimes v)&=(\lambda f+\mu g)(u,v)=\lambda f(u,v)+\mu g(u,v)\\&=\lambda\sigma(f)(u\otimes v)+\mu\sigma(g)(u\otimes v)\\&=(\lambda\sigma(f)+\mu\sigma(g))(u\otimes v).\end{aligned}$$

$\Rightarrow \sigma(\lambda f+\mu g)=\lambda\sigma(f)+\mu\sigma(g)$.

② σ 是满射（显然），单射（$\sigma(f)=0\Rightarrow \forall u,v,\ f(u,v)=0\Rightarrow f=0$）．

即 $\sigma:\mathcal{L}(U,V;W)\to\mathcal{L}(U\otimes V,W)$ 是自然同构． □．

▸ 张量积的基与维数．

- 先来研究 $U\otimes V$ 中等于零的那些张量（仅考虑有限维的情形）

- $u\otimes 0=0\otimes v=0$． (8-28)

证 $u\otimes 0=u\otimes(0+0)=u\otimes 0+u\otimes 0\Rightarrow u\otimes 0=0$．

$0\otimes v=0$ 同理． □

- $\forall u_1,\cdots,u_m\in U$ 线性无关，若 $v_1,\cdots,v_m\in V$ s.t. $\sum\limits_{i=1}^{m}u_i\otimes v_i=0$．

则 $v_1=\cdots=v_m=0$． (8-29)

证 取 $\varphi\in U'$，$\psi\in V'$

$$\begin{tikzcd} U\times V \arrow[r, "\tau"] \arrow[dr, "f(\varphi,\psi)"'] & U\otimes V \arrow[d, dashed, "\rho"] \\ & \mathbb{F} \end{tikzcd}$$

考虑 $f(\varphi,\psi)\in\mathcal{L}(U,V;\mathbb{F})$

s.t. $f(\varphi,\psi)(u,v)=\varphi(u)\psi(v)$

$\Rightarrow\exists$ 唯一的 $\rho\in\mathcal{L}(U\otimes V,\mathbb{F})$ s.t. $\rho(u\otimes v)=\varphi(u)\psi(v)$．

$\sum_{i=1}^{m} u_i \otimes v_i = 0 \Rightarrow \sum_{i=1}^{m} \rho(u_i \otimes v_i) = \sum_{i=1}^{m} \varphi(u_i)\psi(v_i) = 0$.

上式是对 $\forall \varphi \in U', \psi \in V'$ 均成立的

取 $u_1, \cdots, u_m$ 的对偶组 $\varphi_1, \cdots, \varphi_m$ s.t. $\varphi_i(u_j) = \delta_{i,j}$.

$\Rightarrow 0 = \sum_{i=1}^{m} \varphi_k(u_i)\psi(v_i) = \psi(v_k) \Rightarrow v_k = 0 \quad (k = 1, \cdots, m)$.

即 $v_1 = \cdots = v_m = 0$. □

※ $u \otimes v = 0$ 当且仅当 $u = 0$ 或 $v = 0$. (8-30)

· 设 $u_1, \cdots, u_m$ 是 U 的基，$v_1, \cdots, v_n$ 是 V 的基，则 $U \otimes V$ 有基：

$\{u_i \otimes v_j \mid i = 1, \cdots, m;\ j = 1, \cdots, n\}$. (8-31)

证 i) 张成：$\forall u \otimes v \in U \otimes V$，$u = \sum_{i=1}^{m} a_i u_i$. $v = \sum_{j=1}^{n} b_j v_j$.

$\Rightarrow u \otimes v = \sum_{i=1}^{m} a_i u_i \otimes \sum_{j=1}^{n} b_j v_j = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j (u_i \otimes v_j)$

即所有 $u \otimes v$ 的线性组合均可表为 $u_i \otimes v_j$ 的线性组合 .

ii) 线性无关：$\exists \lambda_{i,j} \in \mathbb{F}$ s.t. $\sum_{i=1}^{m}\sum_{j=1}^{n} \lambda_{i,j}(u_i \otimes v_j) = 0$.

$\Rightarrow \sum_{i=1}^{m} u_i \otimes \left(\sum_{j=1}^{n} \lambda_{i,j} v_j\right) = 0$

由定理(8-29)得 $\sum_{j=1}^{n} \lambda_{ij} v_j = 0 \quad \forall i = 1, \cdots, m$.

$\Rightarrow \lambda_{i,j} = 0, \ \forall i = 1, \cdots, m;\ j = 1, \cdots, n$.

各 $u_i \otimes v_j$ 线性无关 . □

※ $\dim(U \otimes V) = \dim U \cdot \dim V$ (8-32)

· 利用张量积的基，有张量的唯一表示：

$\forall \alpha \in U \otimes V, \ \alpha = \sum_{i=1}^{m}\sum_{j=1}^{n} \lambda_{i,j}(u_i \otimes v_j)$. (8-33)

· 结合(8-33)与(8-29)，可以得到张量的其它唯一表示：

i) $\forall \alpha \in U \otimes V$，若 $u_1, \cdots, u_m$ 是 U 的基，

则 $\exists$ 唯一的 $y_1, \cdots, y_m \in V$ s.t. $\alpha = \sum_{i=1}^{m} u_i \otimes y_i$ (8-34)

ii) $\forall \alpha \in U \otimes V$，若 $v_1, \cdots, v_n$ 是 V 的基 .

则 $\exists$ 唯一的 $x_1, \cdots, x_n \in U$ s.t. $\alpha = \sum_{i=1}^{n} x_i \otimes v_i$ (8-35)

· 任意非零张量 $\alpha \in U \otimes V$ 都能唯一写成 $\alpha = \sum_{i=1}^{p} x_i \otimes y_i$.

其中$x_1,\cdots,x_p$与$y_1,\cdots,y_p$分别是U与V中的线性无关组. (8-36)

▸ 张量积的交换律与结合律.

· 有自然同构：$(U\otimes V)\otimes W\cong U\otimes(V\otimes W)$. (8-37)

$\sigma:(u\otimes v)\otimes w\mapsto u\otimes(v\otimes w)$.

$$\begin{tikzcd}(U\otimes V)\times W \arrow[r, "\tau"] \arrow[dr, "f"'] & (U\otimes V)\otimes W \arrow[d, dashed, "\sigma"] \\ & U\otimes(V\otimes W)\end{tikzcd}$$

证 $\dim(U\otimes V)\otimes W=\dim U\cdot\dim V\cdot\dim W$
$=\dim U\otimes(V\otimes W)$.

易知σ是$(U\otimes V)\otimes W$到$U\otimes(V\otimes W)$的自然同构. □

※ 同构视为等同，有$(U\otimes V)\otimes W=U\otimes(V\otimes W)$

记为$U\otimes V\otimes W$，即U,V,W的张量积.

+ n个向量空间的张量积仍用万有性定义，只是把双线性换成多线性，在此不再赘述.

· 有自然同构：$U\otimes V\cong V\otimes U$. $\sigma:u\otimes v\mapsto v\otimes u$. (8-38)

证 $\dim(U\otimes V)=\dim U\cdot\dim V=\dim(V\otimes U)$.

$$\begin{tikzcd}U\times V \arrow[r, "\tau"] \arrow[dr, "f"'] & U\otimes V \arrow[d, dashed, "\sigma"] \\ & V\otimes U\end{tikzcd}$$

易知σ是$U\otimes V$到$V\otimes U$的自然同构. □

※ 张量积有"交换律". 但该自然同构一般不视为等同.

若$U\cap V\neq\varnothing$，则有$(U\otimes V)\cap(V\otimes U)\neq\varnothing$.

考虑$\alpha\in(U\otimes V)\cap(V\otimes U)$，令$\alpha=\sigma(\alpha)$会造成混乱.

※ 两个向量空间可以视为等同，若：

i) 存在不依赖基的选取的同构映射

ii) 交集为空.

8.3 线性映射的张量积

· 本节的中心问题是$\mathcal{L}(U_1,U_2)\otimes\mathcal{L}(V_1,V_2)\cong\mathcal{L}(U_1\otimes V_1,U_2\otimes V_2)$.

以及两个特例：$\mathcal{L}(U)\otimes\mathcal{L}(V)\cong\mathcal{L}(U\otimes V)$

与$U'\otimes V'\cong(U\otimes V)'\cong\mathcal{L}(U,V;\mathbb{F})$

▸ 一般线性映射的情况

· $\forall \varphi \in \mathcal{L}(U_1, U_2), \psi \in \mathcal{L}(V_1, V_2)$. $\exists$ 唯一的线性映射:

$T(\varphi, \psi): U_1 \otimes V_1 \to U_2 \otimes V_2$. s.t. $\forall u \in U_1, v \in V_1$,

$T(\varphi, \psi)(u \otimes v) = \varphi(u) \otimes \psi(v)$. (8-39)

证 设双线性映射 $f(\varphi, \psi): U_1 \times V_1 \to U_2 \otimes V_2$.

$f(\varphi, \psi)(u, v) = \varphi(u) \otimes \psi(v)$.

$$\begin{tikzcd} U_1 \times V_1 \arrow[r, "\tau"] \arrow[dr, "f(\varphi,\psi)"'] & U_1 \otimes V_1 \arrow[d, dashed, "T(\varphi,\psi)"] \\ & U_2 \otimes V_2 \end{tikzcd}$$

由张量积万有性，有唯一的线性映射 $T(\varphi, \psi)$

s.t. $T(\varphi, \psi)(u \otimes v) = \varphi(u) \otimes \psi(v)$ □

※ $T(\varphi, \psi) \in \mathcal{L}(U_1 \otimes V_1, U_2 \otimes V_2)$.

于是有映射 $T: \mathcal{L}(U_1, U_2) \times \mathcal{L}(V_1, V_2) \to \mathcal{L}(U_1 \otimes V_1, U_2 \otimes V_2)$.

· 以上定义的T的双线性映射. (8-40)

证 $T(\lambda_1 \varphi_1 + \lambda_2 \varphi_2, \psi)(u, v) = (\lambda_1 \varphi_1 + \lambda_2 \varphi_2)(u) \otimes \psi(v)$

$= \lambda_1 (\varphi_1(u) \otimes \psi(v)) + \lambda_2 (\varphi_2(u) \otimes \psi(v))$

$= \lambda_1 T(\varphi_1, \psi)(u, v) + \lambda_2 T(\varphi_2, \psi)(u, v)$

$\Rightarrow T(\lambda_1 \varphi_1 + \lambda_2 \varphi_2, \psi) = \lambda_1 T(\varphi_1, \psi) + \lambda_2 T(\varphi_2, \psi)$. 第二个位置同理. □

※ 由张量积的万有性，$\exists$ 唯一的线性映射 $\sigma: \mathcal{L}(U_1, U_2) \otimes \mathcal{L}(V_1, V_2)$

$\to \mathcal{L}(U_1 \otimes V_1, U_2 \otimes V_2)$, $\sigma(\varphi \otimes \psi)(u \otimes v) = \varphi(u) \otimes \psi(v)$.

易证 σ 是自然同构，即 $\mathcal{L}(U_1, U_2) \otimes \mathcal{L}(V_1, V_2) \cong \mathcal{L}(U_1 \otimes V_1, U_2 \otimes V_2)$.

(这种情况只要维数相同基本都有自然同构 ...)

同构视为等同，线性映射的张量积等于张量积的线性映射:

· $\mathcal{L}(U_1, U_2) \otimes \mathcal{L}(V_1, V_2) = \mathcal{L}(U_1 \otimes V_1, U_2 \otimes V_2)$.

$(\varphi \otimes \psi)(u \otimes v) = \varphi(u) \otimes \psi(v)$. (8-41)

▸ 线性算子的情况

· $\forall \varphi_1, \varphi_2 \in \mathcal{L}(U), \psi_1, \psi_2 \in \mathcal{L}(V)$, 有

$(\varphi_1 \otimes \psi_1)(\varphi_2 \otimes \psi_2) = (\varphi_1 \psi_1) \otimes (\varphi_2 \psi_2)$. (8-42)

证 采用8-41的等同.

$\forall u\in U, v\in V$, $(\varphi_1\otimes\psi_1)(\varphi_2\otimes\psi_2)(u\otimes v)=(\varphi_1\otimes\psi_1)(\varphi_2(u)\otimes\psi_2(v))$

$=\varphi_1\varphi_2(u)\otimes\psi_1\psi_2(v)=(\varphi_1\varphi_2\otimes\psi_1\psi_2)(u\otimes v)$

$\Rightarrow(\varphi_1\otimes\psi_1)(\varphi_2\otimes\psi_2)=\varphi_1\varphi_2\otimes\psi_1\psi_2$. □

· $\varphi_1\otimes\varphi_2$可逆当且仅当$\varphi_1$与$\varphi_2$都可逆,且$(\varphi_1\otimes\varphi_2)^{-1}=\varphi_1^{-1}\otimes\varphi_2^{-1}$. (8-43)

证 充分性:若φ_1,φ_2可逆

$(\varphi_1\otimes\varphi_2)(\varphi_1^{-1}\otimes\varphi_2^{-1})=I\otimes I=id$

$\Rightarrow\varphi_1\otimes\varphi_2$可逆, $(\varphi_1\otimes\varphi_2)^{-1}=\varphi_1^{-1}\otimes\varphi_2^{-1}$

必要性:若$\varphi_1\otimes\varphi_2$可逆,设$(\varphi_1\otimes\varphi_2)^{-1}=\psi_1\otimes\psi_2$.

设$U\otimes V$有基$\{u_i\otimes v_j\mid i=1,\cdots,m;\ j=1,\cdots,n\}$

$\Rightarrow(\varphi_1\otimes\varphi_2)(\psi_1\otimes\psi_2)(u_i\otimes v_j)=\varphi_1\psi_1(u_i)\otimes\varphi_2\psi_2(v_j)=u_i\otimes v_j$

$\Rightarrow\varphi_1\psi_1(u_i)=u_i$; $\varphi_2\psi_2(v_j)=v_j$

$\Rightarrow\varphi_1\psi_1=I$, $\varphi_2\psi_2=I\Rightarrow(\varphi_1\otimes\varphi_2)^{-1}=\varphi_1^{-1}\otimes\varphi_2^{-1}$ □

· 张量积的零空间与值域

$\forall\varphi\in\mathcal{L}(U),\psi\in\mathcal{L}(V)$,有: (8-44)

i) $Ran(\varphi\otimes\psi)=Ran\,\varphi\otimes Ran\,\psi$

ii) $Null(\varphi\otimes\psi)=Null\,\varphi\otimes V+U\otimes Null\,\psi$

证 i) $u\otimes v\in Ran(\varphi\otimes\psi)$

$\Leftrightarrow\exists\, u_1\otimes v_1\in U\otimes V$ s.t. $u\otimes v=(\varphi\otimes\psi)(u_1\otimes v_1)$

$\Leftrightarrow\exists\, u_1\in U, v_1\in V$ s.t. $u\otimes v=\varphi(u_1)\otimes\psi(v_1)$

$\Leftrightarrow u\otimes v\in Ran\,\varphi\otimes Ran\,\psi$

即$Ran(\varphi\otimes\psi)=Ran\,\varphi\otimes Ran\,\psi$

ii) $\forall u\otimes v\in Null(\varphi\otimes\psi)$

$\Rightarrow(\varphi\otimes\psi)(u\otimes v)=0\Leftrightarrow\varphi(u)\otimes\psi(v)=0$

$\Rightarrow\varphi(u)=0$ 或 $\psi(v)=0$.

$\Rightarrow u\otimes v \in \mathrm{Null}\,\varphi \otimes V$ 或 $U\otimes \mathrm{Null}\,\psi$

$\Rightarrow u\otimes v \in \mathrm{Null}\,\varphi \otimes V + U\otimes \mathrm{Null}\,\psi$.

即 $\mathrm{Null}(\varphi\otimes\psi) \subset \mathrm{Null}\,\varphi\otimes V + U\otimes \mathrm{Null}\,\psi$.

另一方面，$\forall u_1 \in \mathrm{Null}\,\varphi,\ u_2 \in U,\ v_1 \in V,\ v_2 \in \mathrm{Null}\,\psi$

$u_1\otimes v_1 + u_2\otimes v_2 \in \mathrm{Null}\,\varphi\otimes V + U\otimes \mathrm{Null}\,\psi$.

$\Rightarrow (\varphi\otimes\psi)(u_1\otimes v_1 + u_2\otimes v_2)$

$= \varphi(u_1)\otimes\psi(v_1) + \varphi(u_2)\otimes\psi(v_2)$

$= 0\otimes\psi(v_1) + \varphi(u_2)\otimes 0 = 0$

$\Rightarrow u_1\otimes v_1 + u_2\otimes v_2 \in \mathrm{Null}(\varphi\otimes\psi)$

即 $\mathrm{Null}\,\varphi\otimes V + U\otimes\mathrm{Null}\,\psi \subset \mathrm{Null}(\varphi\otimes\psi)$.

最终有 $\mathrm{Null}(\varphi\otimes\psi) = \mathrm{Null}\,\varphi\otimes V + U\otimes \mathrm{Null}\,\psi$. □

▸ 矩阵的张量积.

· $\forall A \in \mathbb{F}^{m,n}$，$B\in\mathbb{F}^{m',n'}$，$A$ 与 B 的 Kronecker 积定义为

$$\{a_{i,j}B\} = \begin{pmatrix} a_{1,1}B & \cdots & a_{1,n}B \\ \vdots & & \vdots \\ a_{m,1}B & & a_{m,n}B \end{pmatrix} \in \mathbb{F}^{mm',nn'}$$

这是矩阵 A 与 B 的张量积. $A\otimes B = \{a_{ij}B\}$.

$$\mathbb{F}^{m,n}\otimes\mathbb{F}^{m',n'} \cong \mathbb{F}^{mm',nn'} \qquad (8\text{-}45)$$

证 $\dim(\mathbb{F}^{m,n}\otimes\mathbb{F}^{m',n'}) = mm'nn'$

$= \dim(\mathbb{F}^{mm',nn'})$.

$$\begin{tikzcd} \mathbb{F}^{m,n}\times\mathbb{F}^{m',n'} \arrow[r, "\tau"'] \arrow[dr, "f"'] & \mathbb{F}^{m,n}\otimes\mathbb{F}^{m',n'} \arrow[d, dashed, "\sigma"] \\ & \mathbb{F}^{mm',nn'} \end{tikzcd}$$

于是 $\mathbb{F}^{m,n}\otimes\mathbb{F}^{m',n'}$ 与 $\mathbb{F}^{mm',nn'}$ 同构.

只要证 $\sigma(A\otimes B) = \{a_{ij}B\}$ 是单射即可（维数相等，从而是同构）.

$\forall A\otimes B \in \mathrm{Null}\,\sigma$，$\{a_{ij}B\} = 0$

$\Rightarrow a_{i,j} = 0, \forall i,j$ 或 $B = 0 \Rightarrow A = 0$ 或 $B = 0$

$\Rightarrow A\otimes B = 0 \Rightarrow \mathrm{Null}\,\sigma = \{0\}$. 即 σ 是单射. □

※ 若 $\varphi \in \mathcal{L}(U_1, U_2)$ 在某些基下的矩阵为 A，$\psi \in \mathcal{L}(V_1, V_2)$ 在某些基下的矩阵为 B，则 $\varphi \otimes \psi \in \mathcal{L}(U_1 \otimes V_1, U_2 \otimes V_2)$ 在对应基下的矩阵为
$A \otimes B = \{a_{ij} B\}$． (8-46)

▶ 对偶空间的情况．

· 考虑 8-41 式 $U_2 = V_2 = \mathbb{F}$ 的情况：
有自然同构 $U' \otimes V' \cong (U \otimes V)'$
同构视为等同，$(\varphi \otimes \psi)(u \otimes v) = \varphi(u)\psi(v)$． (8-47)

※ 对偶的张量积等于张量积的对偶．

※ 结合 8-47 式与 8-27 式：
$U' \otimes V' \cong (U \otimes V)' = \mathcal{L}(U \otimes V, \mathbb{F}) \cong \mathcal{L}(U, V; \mathbb{F})$． (8-48)
同构视为等同，引出张量积的又一等价定义：

· <定义Ⅳ> 设 U, V 是向量空间，U' 与 V' 的张量积 $U' \otimes V'$ 定义为双线性形式全体 $\mathcal{L}(U, V; \mathbb{F})$．
$\forall \varphi \in U', \psi \in V', (\varphi \otimes \psi)(u, v) = \varphi(u)\psi(v)$．

※ 相似地，$U \otimes V$ 可定义为 $\mathcal{L}(U', V'; \mathbb{F})$． (8-49)

8.4 张量

※ (略过 §8.2 与 §8.3 时，请复习张量积的 <定义Ⅳ>)．

▶ Einstein 约定

· Einstein 约定是张量分析／微分几何中常用的记号系统，用于简化书写．

※ ① 对于成对上下出现的指标（称为哑指标），省略求和号而自动求和
向量的坐标展开：$v = \sum_{i=1}^{n} x_i v_i \longrightarrow v = x^i v_i$．

※ ② 向量的基指标在下，坐标分量指标在上．

※ ③ 矩阵 $A = \{a_{i,j}\} \longrightarrow A = \{a^i_j\}$，行指标在上，列指标在下．

※ ④ Kronecker 符号 $\delta_{i,j} \longrightarrow \delta^i_j$

矩阵作用于向量：$(\vec{y}=A\vec{x})$ $y_i=\sum_{i=1}^{n}a_{i,j}x_j \longrightarrow y^i=a^i_j x^j$

※ ⑤ 求和后，未消去的指标保持上下位置不变.

矩阵乘法：$(C=AB)$ $c_{i,j}=\sum_{k=1}^{n}a_{i,k}b_{k,j} \longrightarrow c^i_j=a^i_k b^k_j$

※ ⑥ 对偶基的向量（称为余向量）指标在上.

$v_1,\cdots,v_n$ 的对偶基为 $\varphi^1,\cdots,\varphi^n$，有 $\varphi^i(v_j)=\delta^i_j$.

· 协变向量 v. 逆变向量.

考虑 V 上有两组基 $v_1,\cdots,v_n$ 和 $\tilde{v}_1,\cdots,\tilde{v}_n$，对应的对偶基分别是 $\varphi^1,\cdots,\varphi^n$ 和 $\tilde{\varphi}^1,\cdots,\tilde{\varphi}^n$.

若 $v_1,\cdots,v_n$ 到 $\tilde{v}_1,\cdots,\tilde{v}_n$ 的过渡矩阵为 $T=\{t^i_j\}$，即 $\tilde{v}_j=t^i_j v_i$

则 $\varphi^1,\cdots,\varphi^n$ 到 $\tilde{\varphi}^1,\cdots,\tilde{\varphi}^n$ 的过渡矩阵为 $(T^{-1})^T$ （8-6式）.

即 $\tilde{\varphi}^j=(t^{-1})^j_i\varphi^i$ 或 $\varphi^i=t^i_j\tilde{\varphi}^j$ (8-50)

注意到对偶基的基变换"方向"恰与原向量空间的基变换相反.

（以原空间基变换方向为准）基向量称为协变向量，对偶基向量称为逆变向量

· 协变坐标 v. 逆变坐标.

记号如前述，$\forall v\in V$，$v=x^i v_i=\tilde{x}^i\tilde{v}_i$

在坐标变换中，坐标分量 $(x^1,\cdots,x^n)^T$ 与 $(\tilde{x}^1,\cdots,\tilde{x}^n)^T$ 满足：

$x^i=t^i_j\tilde{x}_j$，与对偶基变换"方向"相同.

因此原向量空间的坐标称为逆变坐标，对偶空间的坐标称为协变坐标.

※ ⑦ 协变量的指标在下，逆变量的指标在上.

▸ 张量积 $T^p_q(V)$.

· 本节的核心是研究一个向量空间上的张量积.

· 设 V 是有限维向量空间，张量积 $T^p_q(V)=\underbrace{V\otimes\cdots\otimes V}_{p\text{个}V}\otimes\underbrace{V'\otimes\cdots\otimes V'}_{q\text{个}V'}$

称为 V 上的 (p, q) 型张量积，其中的元素称为 (p, q) 型张量，或

p次逆变，q次协变张量.

- $T^p(V) = \underbrace{V \otimes \cdots \otimes V}_{p\text{个}V}$ 称为V上的p次逆变张量积；

  $T_q(V) = \underbrace{V' \otimes \cdots \otimes V'}_{q\text{个}V'}$ 称为V上的q次协变张量积.

※ 根据张量积<定义Ⅳ>，张量积$T^p_q(V)$等同于多线性映射全体：
$\mathcal{L}(\underbrace{V', \cdots, V'}_{q\text{个}V'}; \underbrace{V, \cdots, V}_{p\text{个}V}; \mathbb{F})$. 有人把$(p, q)$型张量比作一台有$p$个上槽和$q$个下槽的机器. 在上槽输入线性泛函，下槽输入向量，便输出一个数*.

- 一些张量的例子：

  $v \in V$是V上的$(1,0)$型张量，$V = T^1_0(V)$；

  $\varphi \in V'$是V上的$(0,1)$型张量，$V' = T^0_1(V)$；

  双线性形式是V上的$(0,2)$型张量，$\mathcal{L}(V, V; \mathbb{F}) = T^0_2(V)$.

- 设V有基$v_1, \cdots, v_n$，对偶基$\varphi^1, \cdots, \varphi^n$，则$T^p_q(V)$有基：
  $\{v_{i_1} \otimes \cdots \otimes v_{i_p} \otimes \varphi^{j_1} \otimes \cdots \otimes \varphi^{j_q}\}$，张量$t \in T^p_q(V)$可展开为：

  $$t = \xi^{i_1 \cdots i_p}_{j_1 \cdots j_q} v_{i_1} \otimes \cdots \otimes v_{i_p} \otimes \varphi^{j_1} \otimes \cdots \otimes \varphi^{j_q}. \qquad (8\text{-}51)$$

  $\{\xi^{i_1 \cdots i_p}_{j_1 \cdots j_q}\}$称为张量$t$的坐标（有些教材直接把$\xi$称为张量）.

  $i_1, \cdots i_p$称为逆变指标，$j_1, \cdots, j_q$称为协变指标.

▶ 张量的运算

- 张量积空间$T^p_q(V)$也是向量空间，向量的加法对张量也成立：

  $$(a+b)^{i_1 \cdots i_p}_{j_1 \cdots j_q} = a^{i_1 \cdots i_p}_{j_1 \cdots j_q} + b^{i_1 \cdots i_p}_{j_1 \cdots j_q}. \qquad (8\text{-}52)$$

- 张量的乘法就是张量的张量积.

  对(p, q)型张量和(r, s)型张量：

  $t_1 = v_1 \otimes \cdots \otimes v_p \otimes \varphi^1 \otimes \cdots \otimes \varphi^q \in T^p_q(V)$.

  $t_2 = u_1 \otimes \cdots \otimes u_r \otimes \psi^1 \otimes \cdots \otimes \psi^s \in T^r_s(V)$.

Campus

* 梁灿彬，周彬，《微分几何入门与广义相对论》(2nd ed.) P21

有 $t_1\otimes t_2 = v_1\otimes\cdots\otimes v_p\otimes u_1\otimes\cdots\otimes u_r\otimes\varphi^1\otimes\cdots\otimes\varphi^q\otimes\psi^1\otimes\cdots\otimes\psi^s$
$\in T^{p+r}_{q+s}(V)$ (8-53)

※ 注意 $T^1_1(V)\otimes T^1_1(V) = V\otimes V'\otimes V\otimes V'$；$T^2_2(V) = V\otimes V\otimes V'\otimes V'$

严格说 $T^1_1(V)\otimes T^1_1(V)\cong T^2_2(V)$，把同构视为等同.

$T^p_q(V)\otimes T^r_s(V)\cong T^{p+r}_{q+s}(V)$

- 若 t_1 的坐标为 $\{\xi^{i_1\cdots i_p}_{j_1\cdots j_q}\}$，$t_2$ 的坐标为 $\{\eta^{k_1\cdots k_r}_{l_1\cdots l_s}\}$，
  则 $t_1\otimes t_2$ 的坐标 $\{\zeta^{i_1\cdots i_p k_1\cdots k_r}_{j_1\cdots j_q l_1\cdots l_s}\} = \{\xi^{i_1\cdots i_p}_{j_1\cdots j_q}\eta^{k_1\cdots k_r}_{l_1\cdots l_s}\}$ (8-54)

※ 张量乘积的坐标等于张量坐标的乘积.

▸ 张量的基变换

- 设 V 上基 $v_1,\cdots,v_n$ 到 $\tilde v_1,\cdots,\tilde v_n$ 的过渡矩阵 $T=\{t^i_j\}$
  对偶基 $\varphi^1,\cdots,\varphi^n$ 到 $\tilde\varphi^1,\cdots,\tilde\varphi^n$ 的过渡矩阵 $(T^{-1})^T=\{s^i_j\}$
  易证关系 $t^i_r s^r_j = \delta^i_j$ (8-55)
- (p,q) 型张量 $t = \xi^{i_1\cdots i_p}_{j_1\cdots j_q} v_{i_1}\otimes\cdots\otimes v_{i_p}\otimes\varphi^{j_1}\otimes\cdots\otimes\varphi^{j_q}$
  $= \tilde\xi^{k_1\cdots k_p}_{l_1\cdots l_q}\tilde v_{k_1}\otimes\cdots\otimes\tilde v_{k_p}\otimes\tilde\varphi^{l_1}\otimes\cdots\otimes\tilde\varphi^{l_q}$

坐标变换为 $\xi^{i_1\cdots i_p}_{j_1\cdots j_q} = t^{i_1}_{k_1}\cdots t^{i_p}_{k_p} s^{l_1}_{j_1}\cdots s^{l_p}_{j_q}\tilde\xi^{k_1\cdots k_p}_{l_1\cdots l_q}$ (8-56)

或 $\tilde\xi^{k_1\cdots k_p}_{l_1\cdots l_q} = \underbrace{t^{j_1}_{l_1}\cdots t^{j_q}_{l_q}}_{\text{协变指标}}\underbrace{s^{k_1}_{i_1}\cdots s^{k_p}_{i_p}}_{\text{逆变指标}}\xi^{i_1\cdots i_p}_{j_1\cdots j_q}$ (8-57)

▸ 缩并

- 考虑映射 $f: \underbrace{V\times\cdots\times V}_{p\text{个}V}\times\underbrace{V'\times\cdots\times V'}_{q\text{个}V'}\longrightarrow T^{p-1}_{q-1}(V)$

定义为 $f(v_1,\cdots,v_p,\varphi^1,\cdots,\varphi^q) = \varphi^1(v_1)(v_2\otimes\cdots\otimes v_p\otimes\varphi^2\otimes\cdots\otimes\varphi^q)$

显然 f 是多线性的. 由张量积的万有性，存在唯一的线性映射：

$C^1_1: T^p_q(V)\to T^{p-1}_{q-1}(V)$，定义为：

$C^1_1(v_1\otimes\cdots\otimes v_p\otimes\varphi^1\otimes\cdots\otimes\varphi^q) = \varphi^1(v_1)(v_2\otimes\cdots\otimes v_p\otimes\varphi^2\otimes\cdots\otimes\varphi^q)$

C^1_1 称为对协变指标 1 和逆变指标 1 的缩并.

· 若 t 的坐标为 $\{\xi^{i_1\cdots i_p}_{j_1\cdots j_q}\}$，则 $C^1_1(t)$ 坐标为 $\{\eta^{i_2\cdots i_p}_{j_2\cdots j_q}\} = \{\xi^{i_1\cdots i_p}_{j_1\cdots j_q}\delta^{j_1}_{i_1}\}$

※ 类似地，可定义 C^k_l 为对协变指标 l 和逆变指标 k 的缩并：

$$C^k_l(v_1\otimes\cdots\otimes v_p\otimes\varphi^1\otimes\cdots\otimes\varphi^q) = (v_1\otimes\cdots\otimes v_{k-1}\otimes v_{k+1}\otimes\cdots\otimes v_p\otimes\varphi^1\otimes\cdots\otimes\varphi^{l-1}\otimes\varphi^{l+1}\otimes\cdots\otimes\varphi^q)\cdot\varphi^l(v_k)$$

· 张量作用于向量

由于维数相同，有同构 $T^1_1(V)\cong\mathcal{L}(V)$。若要把 $(1,1)$ 型张量作用于 V 中的向量（以线性算子的形式），可以用先张量积后缩并的方式。

即 $\forall v\otimes\varphi\in T^1_1(V)$，$w\in V$，有 $\sigma(v\otimes\varphi)\in\mathcal{L}(V)$：

$\sigma(v\otimes\varphi)(w) := C^2_1(v\otimes w\otimes\varphi) = \varphi(w)v$

· 以上定义的 $\sigma: T^1_1(V)\to\mathcal{L}(V)$ 是自然同构。 (8-58)

证 σ 的线性是显然的，要证同构只要证 σ 是单射：

$\forall v\otimes\varphi\in \mathrm{Null}\,\sigma$，有 $\forall w\in V$，$\varphi(w)v=0$

i) 若 $v=0$，$v\otimes\varphi=0$

ii) 若 $v\neq0$，$\forall w\in V$，$\varphi(w)v=0\Rightarrow\varphi(w)=0\Rightarrow\varphi=0\Rightarrow v\otimes\varphi=0$

即 $\mathrm{Null}\,\sigma=\{0\}$，$\sigma$ 是单射。 □

※ 另一种构造自然同构的方式是通过矩阵：$T^1_1(V)\cong\mathbb{F}^{n,n}\cong\mathcal{L}(V)$。

（只是矩阵涉及到基，不那么"自然"）。

※ 相似地，也可构造自然同构 $T^1_1(V)\cong\mathcal{L}(V')$。

$\forall\psi\in V'$，$\theta(v\otimes\varphi)(\psi) := C^1_2(v\otimes\varphi\otimes\psi) = \psi(v)\varphi$。

· $(1,1)$ 型张量的缩并等于其对应线性算子的迹。 (8-59)

证 $\forall v\otimes\varphi\in T^1_1(V)$，$C^1_1(v\otimes\varphi)=\varphi(v)$

设 $v\otimes\varphi\neq0$，研究 $\sigma(v\otimes\varphi)\in\mathcal{L}(V)$ 的本征值：

$\forall w\in\mathrm{Null}\,\sigma(v\otimes\varphi)\Rightarrow\varphi(w)v=0\Rightarrow\varphi(w)=0\Rightarrow w\in\mathrm{Null}\,\varphi$

$\forall w\in\mathrm{Null}\,\varphi\Rightarrow\varphi(w)v=0\Rightarrow w\in\mathrm{Null}\,\sigma(v\otimes\varphi)$.

$\Rightarrow\mathrm{Null}\,\varphi=\mathrm{Null}\,\sigma(v\otimes\varphi)$

$\Rightarrow \dim \mathrm{Null}\,\sigma(v\otimes\varphi) = \dim \mathrm{Null}\,\varphi = \dim V - 1$

即$\sigma(v\otimes\varphi)$有本征值0，重数为$\dim V-1$

又$\sigma(v\otimes\varphi)(v) = \varphi(v)v \in \mathrm{span}(v)$

即$\sigma(v\otimes\varphi)$有本征值$\varphi(v)$，重数为1

$\Rightarrow \mathrm{tr}\,\sigma(v\otimes\varphi) = (\dim V-1)\cdot 0 + 1\cdot\varphi(v) = \varphi(v) = C_1^1(v\otimes\varphi)$ □

8.5 外积与交错张量

▸ 交错多线性映射

· 多线性映射 $f: V\times\cdots\times V \to W$ 称为交错的，若 $\forall i\neq j\ (v_i = v_j)$
$\Rightarrow f(v_1,\cdots,v_p) = 0$.

· 交错性等价于反对称性.（$\mathrm{char}\,F\neq 2$）. (8-60)

证 i) 若 $f: V\times\cdots\times V\to W$ 是交错的

$0 = f(u+v, u+v, \cdots, v_p)$

$= f(u, u+v, \cdots, v_p) + f(v, u+v, \cdots, v_p)$

$= f(u,u,\cdots,v_p) + f(u,v,\cdots,v_p) + f(v,u,\cdots,v_p) + f(v,v,\cdots,v_p)$

$= f(u,v,\cdots,v_p) + f(v,u,\cdots,v_p)$

$\Rightarrow f(u,v,\cdots,v_p) = -f(v,u,\cdots,v_p)$. 对其余位置同理.

即 f 是反对称的.

ii) 若 $f: V\times\cdots\times V\to W$ 是反对称的

$f(v,v,\cdots,v_p) = -f(v,v,\cdots,v_p)$

$\Rightarrow f(v,v,\cdots,v_p) = 0$. 对其余位置同理.

即 f 是交错的. □

※ 思想与3-1是相同的.

▸ 外积

· 考虑 V 上的 p 次逆变张量积 $T^p(V)$，其中由含平方因子张成的子空间

为 N. 即 $N = \text{span}\{v_1\otimes\cdots\otimes v_p \mid \forall 1\le i\ne j\le p\,(v_i = v_j)\}$

商空间 $\Lambda^p(V) = T^p(V)/N$ 称为 V 的 p 重外积空间.

商映射 $\pi: T^p(V)\to\Lambda^p(V)$, $\pi(v_1\otimes\cdots\otimes v_p) = v_1\wedge\cdots\wedge v_p$.

称为向量 $v_1,\cdots,v_p$ 的外积.

- 外积的基本性质 (8-61)

i) 多线性: $\forall\lambda,\mu\in F$, $i=1,\cdots,p$:

$$v_1\wedge\cdots\wedge(\lambda u_i+\mu v_i)\wedge\cdots\wedge v_p = \lambda(v_1\wedge\cdots\wedge u_i\wedge\cdots\wedge v_p)+\mu(v_1\wedge\cdots\wedge v_i\wedge\cdots\wedge v_p)$$

ii) 交错性: $\forall 1\le i\ne j\le p$,

$$v_i = v_j \Rightarrow v_1\wedge\cdots\wedge v_p = 0.$$

证 i) $v_1\wedge\cdots\wedge v_p = \pi(v_1\otimes\cdots\otimes v_p) = \pi\circ\tau(v_1,\cdots,v_p)$.

π 是线性的, 且 τ 是多线性的, 则 $\pi\circ\tau$ 自然也是多线性的.

ii) $\exists 1\le i\ne j\le p\,(v_i=v_j)\Rightarrow v_1\otimes\cdots\otimes v_p\in N$

即 $v_1\wedge\cdots\wedge v_p = v_1\otimes\cdots\otimes v_p + N = 0+N = 0$. □

- 外积的万有性

$\forall$ 向量空间 W & 交错多线性映射 $f: V\times\cdots\times V\to W$, $\exists$ 唯一 $\varphi\in\mathcal{L}(\Lambda^p(V), W)$ s.t. $\varphi(v_1\wedge\cdots\wedge v_p) = f(v_1,\cdots,v_p)$ 恒成立. (8-62)

$$\begin{tikzcd}
\underbrace{V\times\cdots\times V}_{p\text{个}V} \arrow[rr, "\pi\circ\tau"] \arrow[r, "\tau"] \arrow[dr, "f"'] & T^p(V) \arrow[r, "\pi"] \arrow[d, dashed, "\rho"] & \Lambda^p(V) \arrow[dl, dashed, "\varphi"] \\
& W &
\end{tikzcd}$$

证 由张量积的万有性, 存在唯一的 $\rho\in\mathcal{L}(T^p(V), W)$ s.t. $\forall v_1,\cdots,v_p\in V$,

$\rho(v_1\otimes\cdots\otimes v_p) = f(v_1,\cdots,v_p)$.

则只需证 $\rho(v_1\otimes\cdots\otimes v_p) = \varphi(v_1\wedge\cdots\wedge v_p)$:

$\forall t\in N$, $\rho(t) = 0$ (交错性) 且 $\pi(t) = t+N = 0$.

自然有定义 $\varphi(t) = \rho\circ\pi^{-1}(t) = \rho[t+N] = \rho(t)$. □

※: 记全体交错多线性映射构成的向量空间为 $AL(V,\cdots,V;W)$.

上述证明给出了自然同构 $\mathcal{L}(\Lambda^p(V), W)\cong AL(V,\cdots,V;W)$.

- $\Lambda^p(V)$ 的基

设向量空间 V 有基 $v_1, \cdots, v_n$，由8-31得 $T^p(V)$ 的基：

$\{v_{i_1} \otimes \cdots \otimes v_{i_p} \mid i_m = 1, \cdots, n;\ m = 1, \cdots, p\}$

基在商映射 π 下的像满足反对称性 & 交错性

（显然 π 是满射，则基的像也张成全空间 $\Lambda^p(V)$）.

于是 $\Lambda^p(V)$ 的基：$\{v_{i_1} \wedge \cdots \wedge v_{i_p} \mid 1 \le i_1 < \cdots < i_p \le n\}$. (8-63)

- $\Lambda^p(V)$ 的维数 $\dim \Lambda^p(V) = C_n^p$ (8-64)

- 外代数

$\forall w_1 = v_1 \wedge \cdots \wedge v_p \in \Lambda^p(V)$，$w_2 = u_1 \wedge \cdots \wedge u_q \in \Lambda^q(V)$

定义外积 $w_1 \wedge w_2 = v_1 \wedge \cdots \wedge v_p \wedge u_1 \wedge \cdots \wedge u_q \in \Lambda^{p+q}(V)$.

由于外积是由商空间定义的，这里要证明上述定义的合理性.

即要证 $\forall t_1 \in T^p(V)$，$t_2 \in T^q(V)$，有 $\pi(t_1 \otimes t_2) = \pi(t_1) \wedge \pi(t_2)$：

设 $t_1 = v_1 \otimes \cdots \otimes v_p + s_1$，$t_2 = u_1 \otimes \cdots \otimes u_q + s_2$，其中 $s_1, s_2 \in N$.

则 $t_1 \otimes t_2 = v_1 \otimes \cdots \otimes v_p \otimes u_1 \otimes \cdots \otimes u_q + s_3$

其中 $s_3 = s_1 \otimes s_2 + v_1 \otimes \cdots \otimes v_p \otimes s_2 + s_1 \otimes u_1 \otimes \cdots \otimes u_q \in N$.

于是 $\pi(t_1 \otimes t_2) = v_1 \wedge \cdots \wedge v_p \wedge u_1 \wedge \cdots \wedge u_q = \pi(t_1) \wedge \pi(t_2)$. □

对 n 维向量空间 V，考虑 $\Lambda^0(V), \Lambda^1(V), \cdots, \Lambda^n(V)$ 的外直和（积空间）

$$\Lambda(V) = \bigoplus_{p=0}^{n} \Lambda^p(V).$$

（注意 $\Lambda^0(V) = \mathbb{F}$，$\Lambda^1(V) = V$，$\Lambda^{n+1}(V) = \Lambda^{n+2}(V) = \cdots = \{0\}$）.

$\Lambda(V)$ 是向量空间，又由外积的定义也是环，于是 $\Lambda(V)$ 是 $\mathbb{F}$ 上的代数

称为 Grassmann 代数或外代数.

※ 代数的定义：

域 $\mathbb{F}$ 上的向量空间 V 称为 $\mathbb{F}$ 上的代数，若定义 V 上的乘法 $V \times V \to V$

满足：$\forall \lambda_1, \lambda_2 \in \mathbb{F}$，$u_1, u_2, v \in V$：$(\lambda_1 u_1 + \lambda_2 u_2)v = \lambda_1(u_1 v) + \lambda_2(u_2 v)$

$\forall \lambda_1, \lambda_2 \in \mathbb{F}$，$u, v_1, v_2 \in V$：$u(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1(u v_1) + \lambda_2(u v_2)$.

- $\Lambda(V)$ 的维数 $\dim \Lambda(V) = \sum_{p=0}^{n} \dim \Lambda^p(V) = \sum_{p=0}^{n} C_n^p = 2^n$ (8-65)

证 对$(x+1)^n$二项式展开，代入$x=1$，即得$2^n=\sum_{p=0}^{n}C_n^p$. □

例 $\dim\Lambda(\mathbb{R}^3)=8$. 基 $\underbrace{e_1,e_2,e_3}_{\Lambda^1(\mathbb{R}^3)},\underbrace{e_1\wedge e_2,e_1\wedge e_3,e_2\wedge e_3}_{\Lambda^2(\mathbb{R}^3)},\underbrace{e_1\wedge e_2\wedge e_3}_{\Lambda^3(\mathbb{R}^3)},\underbrace{1}_{\mathbb{R}}$.

- 交错张量

※ 以上讨论从商空间的角度刻画了交错多线性映射，接下来考虑张量积的一个子空间，即交错张量，从而刻画外积.

- V上的p次协变张量积$T_p(V)\cong\mathcal{L}(V,\cdots,V;F)$.

  $\forall t\in T_p(V)$，若$t(v_1,\cdots,v_p)$是交错多线性映射，则t称为交错张量.
  $T_p(V)$中的全体交错张量构成向量空间，记为$AT_p(V)$.

※ 这里插入一些必要的抽象代数结论：

- 集合$\{1,\cdots,n\}$到自身的双射称为置换σ，$\{1,\cdots,n\}$上全体置换构成的群称为n元对称群S_n.
- 若置换δ满足$\delta(i_1)=i_2$，$\delta(i_2)=i_1$，且其余元素不变，则δ称为一个对换.（即对换元素i_1与i_2）可以证明$\forall\sigma\in S_n$，σ可表示为有限个对换的复合.
- $\forall\sigma\in S_n$，$\exists$对换$\delta_1,\cdots,\delta_k$ s.t. $\sigma=\delta_1\cdots\delta_k$. (8-66)

  若k为奇数，σ称为奇置换；若k为偶数，σ称为偶置换.
- 定义 $sg(\sigma)=\begin{cases}1, & \sigma\text{为偶置换}\\ -1, & \sigma\text{为奇置换}\end{cases}$ $\iff sg(\sigma)=(-1)^k$. (8-67)

  $\forall\sigma,\tau\in S_n$，$sg(\tau\sigma)=sg(\tau)sg(\sigma)$；$sg(\sigma)=sg(\tau)sg(\tau\sigma)$ (8-68)
- 置换作用于张量

  $\forall\sigma\in S_p$，定义$\sigma:T_p(V)\to T_p(V)$如下：

  $\forall t\in T_p(V)$，$v_1,\cdots,v_p\in V$，有$\sigma(t)(v_1,\cdots,v_p)=t(v_{\sigma(1)},\cdots,v_{\sigma(p)})$

※ 易证$\forall\varphi^1\otimes\cdots\otimes\varphi^p\in T_p(V)$，$\sigma(\varphi^1\otimes\cdots\otimes\varphi^p)=\varphi^{\sigma(1)}\otimes\cdots\otimes\varphi^{\sigma(p)}$ (8-69)

即置换对张量的作用是置换指标.

※ 置换可以给出交错张量的等价定义：

· 下列表述等价： (8-70)

i) $t \in T_p(V)$ 是交错张量

ii) $\sigma(t) = sg(\sigma)t$

iii) $\delta(t) = -t$ ($\delta \in S_p$ 是对换).

证 iii) 实质上是反对称性，已证明与 i) 的交错性等价

ii) $\Rightarrow$ iii) 显然.

iii) $\Rightarrow$ ii) $\forall \sigma \in S_p$, $\exists \delta_1, \cdots, \delta_k \in S_n$ s.t. $\sigma = \delta_1 \cdots \delta_k$, $sg(\sigma) = (-1)^k$.

$\Rightarrow \sigma(t) = \delta_1 \cdots \delta_k(t) = (-1)^k t = sg(\sigma)t$. □

· 交错化算子

$\forall t \in T_p(V)$，定义交错化算子 $\mathcal{A} \in \mathcal{L}(T_p(V))$：

$\mathcal{A}(t) = \sum_{\sigma \in S_p} sg(\sigma) \cdot \sigma(t)$

※ $\forall \varphi^1 \otimes \cdots \otimes \varphi^p \in T_p(V)$, $\mathcal{A}(t) = \sum_{\sigma \in S_p} sg(\sigma) \cdot \sigma(\varphi^1 \otimes \cdots \otimes \varphi^p)$

$= \sum_{\sigma \in S_p} sg(\sigma) \cdot (\varphi^{\sigma(1)} \otimes \cdots \otimes \varphi^{\sigma(p)})$. (8-71)

该形式与行列式的展开十分相似：

$\det A = \sum_{\sigma \in S_n} sg(\sigma)\, a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$.

其关系将随后解释.

· 交错化算子的性质 (8-72)

$\forall t \in T_p(V)$, $\mathcal{A} \in \mathcal{L}(T_p(V))$, $\tau \in S_p$：

i) $\tau(\mathcal{A}t) = \mathcal{A}(\tau(t)) = sg(\tau) \cdot \mathcal{A}(t)$.

ii) $\mathcal{A}(t) \in AT_p(V)$ (交错化算子把张量映成交错张量)

iii) 若 $t \in AT_p(V)$，则 $\mathcal{A}(t) = p!\, t$.

证 i) $\tau(\mathcal{A}(t)) = \tau\left(\sum_{\sigma \in S_p} sg(\sigma) \cdot \sigma(t)\right) = \sum_{\sigma \in S_p} sg(\sigma) \cdot \tau\sigma(t)$

$= sg(\tau)\left(\sum_{\sigma \in S_p} sg(\tau\sigma) \cdot \tau\sigma(t)\right) = sg(\tau) \cdot \mathcal{A}(t)$

$\mathcal{A}(\tau(t)) = \sum_{\sigma \in S_p} sg(\sigma) \cdot \sigma\tau(t) = sg(\tau)\left(\sum_{\sigma \in S_p} sg(\sigma\tau) \cdot \sigma\tau(t)\right)$

$= sg(\tau)\cdot \mathcal{A}(t)$

ii) 由 i) 与 8-70 得证

iii) 由 8-70, $t\in AT_p(V) \Rightarrow \sigma(t) = sg(\sigma)\cdot t$

$\Rightarrow \mathcal{A}(t) = \sum_{\sigma\in S_p} sg(\sigma)\cdot\sigma(t) = \sum_{\sigma\in S_p} sg(\sigma)\cdot sg(\sigma)\cdot t$

$= \sum_{\sigma\in S_p} t = p!\,t$. □

※ 协变张量作为多线性映射：

$\forall \varphi^1\otimes\cdots\otimes\varphi^p\in T_p(V)$, $(\varphi^1\otimes\cdots\otimes\varphi^p)(v_1,\cdots,v_p) = \varphi^1(v_1)\cdots\varphi^p(v_p)$.

而交错协变张量作用如下

- $\forall \varphi^1\otimes\cdots\otimes\varphi^p\in T_p(V)$, $\mathcal{A}(\varphi^1\otimes\cdots\otimes\varphi^p)\in AT_p(V)$：

$$\mathcal{A}(\varphi^1\otimes\cdots\otimes\varphi^p)(v_1,\cdots,v_p) = \begin{vmatrix} \varphi^1(v_1) & \cdots & \varphi^1(v_p) \\ \vdots & & \vdots \\ \varphi^p(v_1) & \cdots & \varphi^p(v_p)\end{vmatrix} = \det\{\varphi^i(v_j)\}_{i,j=1}^{p} \qquad (8\text{-}73)$$

证 $\mathcal{A}(\varphi^1\otimes\cdots\otimes\varphi^p)(v_1,\cdots,v_p)$

$= \sum_{\sigma\in S_p} sg(\sigma)\cdot\sigma(\varphi^1\otimes\cdots\otimes\varphi^p)(v_1,\cdots,v_p)$

$= \sum_{\sigma\in S_p} sg(\sigma)\,(\varphi^{\sigma(1)}\otimes\cdots\otimes\varphi^{\sigma(p)})(v_1,\cdots,v_p)$

$= \sum_{\sigma\in S_p} sg(\sigma)\,\varphi^{\sigma(1)}(v_1)\cdots\varphi^{\sigma(p)}(v_p)$

$= \det\{\varphi^i(v_j)\}_{i,j=1}^{n}$. □

※ 交错协变张量等同于交错多线性映射，作用方式为行列式.

$AT_p(V)\cong AL(V,\cdots,V;\mathbb{F})$.

- $AT_p(V)$ 的基与维数

交错化算子给出了 $T_p(V)$ 到 $AT_p(V)$ 的满射，因此直接研究 $\mathcal{A}$ 对 $T_p(V)$ 的基的作用即可.

设 V 的基 $e_1,\cdots,e_n$，对偶基为 $\xi^1,\cdots,\xi^n$，则 $AT_p(V)$ 有基：

$\{\mathcal{A}(\xi^{i_1}\otimes\cdots\otimes\xi^{i_p}) \mid 1\le i_1<\cdots<i_p\le n\}$. $\qquad (8\text{-}74)$

证 张成空间是显然的：

$\mathcal{A}(t)=sg(\sigma)\cdot\sigma(t)$

$\Rightarrow \forall\sigma\in S_n$, $span(\mathcal{A}(\xi^{i\sigma(1)}\otimes\cdots\otimes\xi^{i\sigma(p)}))=span(\mathcal{A}(\xi^{i_1}\otimes\cdots\otimes\xi^{i_p}))$

$\Rightarrow AT_p(V)=span\{\mathcal{A}(\xi^{j_1}\otimes\cdots\otimes\xi^{j_p})\mid 1\le j_k\le n,\ 1\le k\le p\}$

$=span\{\mathcal{A}(\xi^{i\sigma(1)}\otimes\cdots\otimes\xi^{i\sigma(p)})\mid\sigma\in S_n,\ 1\le i_1\le\cdots\le i_p\le n\}$

$=span\{\mathcal{A}(\xi^{i_1}\otimes\cdots\otimes\xi^{i_p})\mid 1\le i_1<\cdots<i_p\le n\}$

线性无关：$\exists\lambda_{i_1\cdots i_p}\in\mathbb{F}$, $1\le i_1<\cdots<i_p\le n$,

s.t. $\sum\limits_{1\le i_1<\cdots<i_p\le n}\lambda_{i_1\cdots i_p}\mathcal{A}(\xi^{i_1}\otimes\cdots\otimes\xi^{i_p})=0$.

$\Rightarrow\sum\limits_{1\le i_1<\cdots<i_p\le n}\lambda_{i_1\cdots i_p}\mathcal{A}(\xi^{i_1}\otimes\cdots\otimes\xi^{i_p})(e_{j_1},\cdots,e_{j_p})=0$.

注意 $\mathcal{A}(\xi^{i_1}\otimes\cdots\otimes\xi^{i_p})(e_{j_1},\cdots,e_{j_p})\ne0 \iff \{j_1,\cdots,j_p\}$ 是 $\{i_1,\cdots,i_p\}$ 的一个置换

$\Rightarrow 0=\lambda_{j_1,\cdots,j_p}\mathcal{A}(\xi^{j\sigma(1)}\otimes\cdots\otimes\xi^{j\sigma(p)})(e_{j_1},\cdots,e_{j_p})=\lambda_{j_1\cdots j_p}\det\{\xi^{j\sigma(r)}(e_{j_s})\}$

$=\pm\lambda_{j_1,\cdots,j_p}$

由此证明所有 $\lambda_{i_1,\cdots,i_p}=0$，各 $\mathcal{A}(\xi^{i_1}\otimes\cdots\otimes\xi^{i_p})$ 线性无关. □

※ $\dim AT_p(V)=C_n^p$. (8-75)

▸ 交错张量的外积.

• 定义从外积空间到交错张量积的映射. $\theta:\Lambda^p(V')\to AT_p(V)$

$\forall\varphi^1,\cdots,\varphi^p\in V$, $\theta(\varphi^1\wedge\cdots\wedge\varphi^p)=\mathcal{A}(\varphi^1\otimes\cdots\otimes\varphi^p)$.

$\dim\Lambda^p(V')=\dim AT_p(V)=C_n^p$，于是易证 θ 是 $\Lambda^p(V')$ 到 $AT_p(V)$ 的自然同构.

※ 同构视为等同：$\varphi^1\wedge\cdots\wedge\varphi^p=\mathcal{A}(\varphi^1\otimes\cdots\otimes\varphi^p)\in AT_p(V)$.

对逆变张量同样有 $v_1\wedge\cdots\wedge v_p=\mathcal{A}(v_1\otimes\cdots\otimes v_p)\in AT^p(V)$.

上式也常作为外积空间的等价定义.

※ p次交错协变张量又称为p-形式（微分几何中常用）

• $\Lambda^p(V)$与$\Lambda^p(V')$的作用

由同构 $\Lambda^p(V')\cong AT_p(V)\cong AL(V,\cdots,V;\mathbb{F})$

外积视为交错映射：$(\varphi^1\wedge\cdots\wedge\varphi^p)(v_1,\cdots,v_p)=\det\{\varphi^i(v_j)\}$

由外积万有性，又有同构 $\Lambda^p(V')\cong(\Lambda^p(V))'$：

$$(\varphi^1\wedge\cdots\wedge\varphi^p)(v_1\wedge\cdots\wedge v_p)=\det\{\varphi^i(v_j)\}. \qquad (8\text{-}76)$$

和张量积的 $(\varphi^1\otimes\cdots\otimes\varphi^p)(v_1\otimes\cdots\otimes v_p)=\varphi^1(v_1)\cdots\varphi^p(v_p)$ 类似.

- 交错张量的外积.

$\forall\, t_1\in AT^p(V),\ t_2\in AT^q(V)$:

$$t_1\wedge t_2=\frac{1}{p!\,q!}\mathcal{A}(t_1\otimes t_2)\in AT^{p+q}(V). \qquad (8\text{-}77)$$

证 设 V 的基 $e_1,\cdots,e_n$.

$$p!\,t_1=\mathcal{A}(t_1)=\sum_{1\le i_1<\cdots<i_p\le n}\lambda^{i_1\cdots i_p}e_{i_1}\wedge\cdots\wedge e_{i_p}$$

$$q!\,t_2=\mathcal{A}(t_2)=\sum_{1\le j_1<\cdots<j_q\le n}\eta^{j_1\cdots j_q}e_{j_1}\wedge\cdots\wedge e_{j_q}$$

$$\Rightarrow(p!\,t_1)\wedge(q!\,t_2)=\sum_{1\le i_1<\cdots<i_p\le n}\sum_{1\le j_1<\cdots<j_q\le n}\lambda^{i_1\cdots i_p}\eta^{j_1\cdots j_q}e_{i_1}\wedge\cdots\wedge e_{i_p}\wedge e_{j_1}\wedge\cdots\wedge e_{j_q}$$

$$=\sum_{1\le i_1<\cdots<i_p\le n}\sum_{1\le j_1<\cdots<j_q\le n}\lambda^{i_1\cdots i_p}\eta^{j_1\cdots j_q}\mathcal{A}(e_{i_1}\otimes\cdots\otimes e_{i_p}\otimes e_{j_1}\otimes\cdots\otimes e_{j_q})$$

$$=\mathcal{A}\Big(\sum_{1\le i_1<\cdots<i_p\le n}\sum_{1\le j_1<\cdots<j_q\le n}\lambda^{i_1\cdots i_p}\eta^{j_1\cdots j_q}e_{i_1}\otimes\cdots\otimes e_{i_p}\otimes e_{j_1}\otimes\cdots\otimes e_{j_q}\Big)$$

$$=\mathcal{A}\Big(\big(\sum_{1\le i_1<\cdots<i_p\le n}\lambda^{i_1\cdots i_p}e_{i_1}\otimes\cdots\otimes e_{i_p}\big)\otimes\big(\sum_{1\le j_1<\cdots<j_q\le n}\eta^{j_1\cdots j_q}e_{j_1}\otimes\cdots\otimes e_{j_q}\big)\Big).$$

$$=\mathcal{A}(t_1\otimes t_2)$$

$$\Rightarrow t_1\wedge t_2=\frac{1}{p!\,q!}\mathcal{A}(t_1\otimes t_2) \qquad \square$$

※ 注意当 t_1,t_2 为交错张量时，

$\mathcal{A}(t_1)\wedge\mathcal{A}(t_2)=\mathcal{A}(t_1\otimes t_2)$ 与 $\pi(t_1)\wedge\pi(t_2)=\pi(t_1\otimes t_2)$

形式上是统一的.

- 定理8-77给出了构造外代数的又一方法 $AT(V)=\sum_{p=0}^{n}AT^p(V)$.
- 外积的反交换律

$$\forall\, t_1\in AT^p(V),\ t_2\in AT^q(V),\ t_1\wedge t_2=(-1)^{pq}t_2\wedge t_1. \qquad (8\text{-}78)$$

证 $t_1\wedge t_2\in AT^{p+q}(V)\Rightarrow\forall\,\sigma\in S_{p+q},\ \sigma(t_1\wedge t_2)=sg(\sigma)\cdot t_1\wedge t_2$.

取 $\sigma=\begin{pmatrix}1&\cdots&p&p+1&\cdots&p+q\\ q+1&\cdots&q+p&1&\cdots&q\end{pmatrix}\Rightarrow sg(\sigma)=(-1)^{pq}$

$\forall \varphi^1, \cdots, \varphi^{p+q} \in V'$, $(t_1 \wedge t_2)(\varphi^1, \cdots, \varphi^{p+q})$

$= (-1)^{pq}(t_1 \wedge t_2)(\varphi^{\sigma(1)}, \cdots, \varphi^{\sigma(p+q)})$

$= (-1)^{pq} \cdot \frac{1}{p!q!} \mathcal{A}(t_1 \otimes t_2)(\varphi^{\sigma(1)}, \cdots, \varphi^{\sigma(p+q)})$

$= (-1)^{pq} \cdot \frac{1}{p!q!} \sum_{\tau \in S_{p+q}} sg(\tau) t_1(\varphi^{\tau\cdot\sigma(1)}, \cdots, \varphi^{\tau\cdot\sigma(p)}) t_2(\varphi^{\tau\cdot\sigma(p+1)}, \cdots, \varphi^{\tau\cdot\sigma(p+q)})$

$= (-1)^{pq} \cdot \frac{1}{p!q!} \sum_{\tau \in S_{p+q}} sg(\tau) t_2(\varphi^{\tau(1)}, \cdots, \varphi^{\tau(q)}) t_1(\varphi^{\tau(q+1)}, \cdots \varphi^{\tau(p+q)})$

$= (-1)^{pq} \mathcal{A}(t_2 \otimes t_1) = (-1)^{pq} t_2 \wedge t_1$ □

- 外积的结合律 (8-79)

$\forall t_1 \in AT^p(V), t_2 \in AT^q(V), t_3 \in AT^r(V)$, $(t_1 \wedge t_2) \wedge t_3 = t_1 \wedge (t_2 \wedge t_3)$

证 $\forall \varphi^1, \cdots, \varphi^{p+q+r} \in V'$, $((t_1 \wedge t_2) \wedge t_3)(\varphi^1, \cdots, \varphi^{p+q+r})$

$= \mathcal{A}((t_1 \wedge t_2) \otimes t_3)(\varphi^1, \cdots, \varphi^{p+q+r})$

$= \frac{1}{(p+q)!r!} \sum_{\sigma \in S_{p+q+r}} sg(\sigma)(t_1 \wedge t_2)(\varphi^{\sigma(1)}, \cdots, \varphi^{\sigma(p+q)}) t_3(\varphi^{\sigma(p+q+1)}, \cdots, \varphi^{\sigma(p+q+r)})$

$= \frac{1}{(p+q)!p!q!r!} \sum_{\sigma \in S_{p+q+r}} sg(\sigma) \mathcal{A}(t_1 \otimes t_2)(\varphi^{\sigma(1)}, \cdots, \varphi^{\sigma(p+q)}) t_3(\varphi^{\sigma(p+q+1)}, \cdots, \varphi^{\sigma(p+q+r)})$

$= \frac{1}{p!q!r!} \sum_{\sigma \in S_{p+q+r}} sg(\sigma) t_1(\varphi^{\sigma(1)}, \cdots, \varphi^{\sigma(p)}) t_2(\varphi^{\sigma(p+1)}, \cdots, \varphi^{\sigma(p+q)}) t_3(\sim)$.

$= \frac{1}{p!q!r!} \mathcal{A}(t_1 \otimes t_2 \otimes t_3)(\varphi^1, \cdots, \varphi^{p+q+r})$

同理 $t_1 \wedge (t_2 \wedge t_3) = \frac{1}{p!q!r!} \mathcal{A}(t_1 \otimes t_2 \otimes t_3)$. □

※ 外代数是结合代数.

- 行列式
- 向量组的行列式

这是对§3.1行列式定义的Formalization:

设V是n维向量空间，基为 $e_1, \cdots, e_n$，行列式 $det \in \Lambda^n(V') \cong AL(V, \cdots, V; \mathbb{F})$ 是满足 $det(e_1, \cdots, e_n) = 1$ 的 n-形式(多重线性泛函).

- 线性算子的行列式.

传统上 $T: V \to V$ 的行列式定义为T在任意基下矩阵的行列式.

这里给出一个不依赖基的行列式等价定义:

$\forall T \in \mathcal{L}(V)$，$T$ 诱导了同构 $\det T \in \mathcal{L}(\Lambda^n(V'))$：

设 $\omega \in \Lambda^n(V')$，$\forall v_1, \cdots, v_n \in V$，

$$\det T(\omega)(v_1, \cdots, v_n) = \omega(T(v_1), \cdots, T(v_n)). \qquad (8\text{-}80)$$

$\det T$ 作为一维空间中的同构，相当于一个数 $\mathbb{F}$，称为算子 T 的行列式.

证 设 V 的基 $e_1, \cdots, e_n$，对偶基 $\xi^1, \cdots, \xi^n$.

设 $\omega = k\xi^1 \wedge \cdots \wedge \xi^n$，$v_1 \wedge \cdots \wedge v_n = \lambda e_1 \wedge \cdots \wedge e_n$，$T(e_j) = a_j^i e_i$

$\omega(T(v_1), \cdots, T(v_n))$

$$
\begin{aligned}
&= k(\xi^1 \wedge \cdots \wedge \xi^n)(T(v_1) \wedge \cdots \wedge T(v_n)) \\
&= k\lambda(\xi^1 \wedge \cdots \wedge \xi^n)(T(e_1) \wedge \cdots \wedge T(e_n)) \\
&= \sum_{i_1, \cdots, i_n = 1}^{n} k\lambda(\xi^1 \wedge \cdots \wedge \xi^n)(a_1^{i_1} e_{i_1} \wedge \cdots \wedge a_n^{i_n} e_{i_n}) \\
&= \sum_{\sigma \in S_n} k\lambda(\xi^1 \wedge \cdots \wedge \xi^n)(a_1^{\sigma(1)} e_{\sigma(1)} \wedge \cdots \wedge a_n^{\sigma(n)} e_{\sigma(n)}) \\
&= k\lambda \sum_{\sigma \in S_n} \operatorname{sg}(\sigma)\, a_1^{\sigma(1)} \cdots a_n^{\sigma(n)} (\xi^1 \wedge \cdots \wedge \xi^n)(e_1 \wedge \cdots \wedge e_n) \\
&= k\lambda \det\{a_j^i\}.
\end{aligned}
$$

$\det T(\omega)(v_1, \cdots, v_n)$

$$
\begin{aligned}
&= \det T \cdot k(\xi^1 \wedge \cdots \wedge \xi^n) \cdot \lambda(e_1 \wedge \cdots \wedge e_n) \\
&= k\lambda \det T
\end{aligned}
$$

$\Rightarrow \det T = \det\{a_j^i\}$. □