矩阵论（Matrix Theory）

数理与统计学院
SUES

矩阵论(MATRIX THEORY)

◆ 矩阵的广泛应用

	X_1	X_2	$\cdots$	X_k	$\cdots$	X_p
S_1	x_{11}	x_{12}	$\cdots$	x_{1k}	$\cdots$	x_{1p}
S_2	x_{21}	x_{22}	$\cdots$	x_{2k}	$\cdots$	x_{2p}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
S_n	x_{n1}	x_{n2}	$\cdots$	x_{nk}	$\cdots$	x_{np}

由上表，记

$$\mathbf{X}_{(i)} = (x_{i1}, x_{i2}, \cdots, x_{ip})', \quad i = 1, 2, \cdots, n$$

和

$$\mathbf{X}_j = (x_{1j}, x_{2j}, \cdots, x_{nj})', \quad j = 1, 2, \cdots, p$$

分别表示第i个样品的p个变量的观测值和不同样品的第j个变量X_j的n次观测值. 在具体观测之前，它们分别为p维和n维随机向量.

数据表可以用矩阵形式表示，记为$\mathbf{X}$，即

$$\mathbf{X} = (x_{ij})_{n\times p} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'_{(1)} \\ \mathbf{X}'_{(2)} \\ \vdots \\ \mathbf{X}'_{(n)} \end{pmatrix} = (\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_p).$$

称$\mathbf{X}$为样本数据阵(或样本资料阵). 在具体观测之前，它是一个随机矩阵.

矩阵论(MATRIX THEORY)

◆ 马科维茨的均值-方差模型

哈里.马科维茨(Harry M. Markowitz)于1952年在《财务学杂志》上发表论文《Portfolio Selection》，该论文最早采用风险资产的期望收益率(均值)和用方差(或标准差)代表的风险来研究资产组合和选择问题，堪称现代金融理论史上的里程碑，标志着现代组合投资理论的开端. 1989年美国运筹学会、管理科学协会联合授予马科维茨、冯·诺伊曼运筹学理论奖，以表彰他们在证券组合选择理论、稀疏矩阵技术、SIMSCRIPT程序语言等方面所作的理论突破和技术创新工作.由于其出色的、开创性的工作，马科维茨与威廉夏普及默顿米勒分享了1990年诺贝尔经济学奖.

◆ 马科维茨的均值-方差模型

$$\min \quad x^T \Sigma x$$

$$\text{s.t.} \begin{cases} 1^T x = 1, \\ r^T x \geq r_{\min}, \\ x \geq 0. \end{cases}$$

◆ 主成分分析

主成分分析(Principal Component Analysis，PCA)是将多个变量通过线性变换以选出较少个数重要变量的一种多元统计分析方法，在人口统计学、数量地理学、分子动力学模拟、数学建模、数理分析等学科领域有广泛应用.

PCA首先是由**K. 皮尔森(1901)**对非随机变量引入的，尔后**H. 霍特林**将此方法推广到随机向量的情形.

基本思想：研究如何通过少数几个主成分来揭示多个变量间的内部结构，即从原始变量中导出少数几个主成分，使它们尽可能多地保留原始变量的信息，且彼此间互不相关.数学上的处理就是将原来p个指标作线性组合，作为新的综合指标.

矩阵论(MATRIX THEORY)

◆ 主成分分析

PCA的优化模型(续)

第一主成分的优化模型：

$$\alpha_1 = \underset{\alpha}{\operatorname{argmax}} \ \alpha^T \hat{\Sigma} \alpha, \quad \text{s.t.} \ \|\alpha\|_2 = 1,$$

其中$\hat{\Sigma} = (\mathbf{X}^T\mathbf{X})/(N-1)$为样本协方差矩阵.

其余主成分的优化模型：

$$\alpha_{k+1} = \underset{\alpha}{\operatorname{argmax}} \ \alpha^T \hat{\Sigma} \alpha,$$

$$\text{s.t.} \ \|\alpha\|_2 = 1, \alpha^\top \alpha_l = 0, \forall 1 \le l \le k.$$

矩阵论(MATRIX THEORY)

本课程的主要内容

- 线性空间与线性变换
- 内积空间
- 范数理论
- Jordan标准型
- 矩阵分析
- 矩阵分解
- 矩阵广义逆
- 特殊矩阵
- 矩阵的特殊积等

第一章

线性空间与线性变换

◆ 向量空间

对于平面 R^2 中的任意向量，我们已定义过加法及数乘两种运算，而且这两种运算是封闭的，即运算后的结果仍在 R^2 中.

而且这两种运算满足下面8条运算律：

对$\forall\alpha$、β、$\gamma\in R^2$，$\forall k$、$l\in R$，成立

(1) 加法交换律：$\alpha+\beta=\beta+\alpha$，

(2) 加法结合律：$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$，

(3) 具有**加法单位元（零向量）** $\mathbf{0} \in \boldsymbol{R}^2$，使得

$$\alpha + \mathbf{0} = \alpha$$

(4) 具有**加法逆元（负向量）** $-\alpha \in \boldsymbol{R}^2$，使得

$$\alpha + (-\alpha) = \mathbf{0}$$

(5) 数乘的结合律：$k(l\alpha) = (kl)\alpha$

(6) 数乘的单位元：$1 \cdot \alpha = \alpha$

(7) 分配律1：$k(\alpha + \beta) = k\alpha + k\beta$

(8) 分配律2：$(k + l)\alpha = k\alpha + l\alpha$

由线性代数的知识，二维向量空间 R^2 显然可推广到 n 维向量空间 R^n． 并且数乘所依赖的实数域 R 也可推广到复数域 C． 相应的向量空间分别称为实向量空间和复向量空间.

我们知道，向量是特殊的矩阵． 所有 $m\times n$ 阶的实矩阵的集合 $R^{m\times n}$ 对矩阵的加法和数乘封闭，并且也满足上述8条运算律. 因此也是“实向量空间”.

不过这里的“向量”是实矩阵！

§1.1线性空间的基本概念

线性空间是线性代数最基本的概念之一，是矩阵论中极其重要的概念之一. 它是向量空间在元素和线性运算上的推广和抽象.

线性空间中的元素可以是向量、矩阵、多项式、函数等，线性运算可以是我们熟悉的一般运算，也可以是各种特殊的运算.

1.1.2 线性空间的定义与性质

一、线性空间的概念

定义1.1.4 如果非空集合 V 对于加法及数乘两种运算封闭，并且对于加法和数乘满足下面8条运算律，那么就称集合 V 为数域 P 上的**线性空间**或**向量空间**：

对$\forall \alpha$、β、$\gamma \in V$，$\forall k$、$l \in P = R$ 或$P = C$，成立

(1) 加法交换律：$\alpha + \beta = \beta + \alpha$，

(2) 加法结合律：$(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$，

(3) 具有加法单位元（零向量）$\mathbf{0}\in V$，使得

$$\alpha+\mathbf{0}=\alpha$$

(4) 具有加法逆元（负向量）$-\alpha\in V$，使得

$$\alpha+(-\alpha)=\mathbf{0}$$

(5) 数乘的结合律：$k(l\alpha)=(kl)\alpha$

(6) 数乘的单位元：$1\cdot\alpha=\alpha$

(7) 分配律1：$k(\alpha+\beta)=k\alpha+k\beta$

(8) 分配律2：$(k+l)\alpha=k\alpha+l\alpha$

注意：这里我们不再关心元素的特定属性，而且我们也不用关心这些线性运算（加法和数乘）的具体形式.

例如：在正实数集$R^+=\{a\,|\,a>0, a\in R\}$中定义加法“$\oplus$”和数乘“$\otimes$”运算如下：

$$a\oplus b=ab,\ \ \lambda\otimes a=a^{\lambda},\ \ \forall a,b\in R^+, \lambda\in R$$

则R^+是数域R上的线性空间.

事实上，有

$$\forall a,b \in R^+ \Rightarrow a \oplus b = ab \in R^+;$$

$$\forall \lambda \in R, a \in R^+ \Rightarrow \lambda \otimes a = a^\lambda \in R^+.$$

所以对定义的加法与数乘运算封闭.

下面一一验证八条线性运算规律：

(1) $a \oplus b = ab = ba = b \oplus a;$

(2) $(a \oplus b) \oplus c = (ab) \oplus c = (ab)c = a(bc) = a \oplus (b \oplus c);$

(3) R^+中存在零元素 1, 对于$\forall a \in R^+$, 有

$$a \oplus 1 = a \cdot 1 = a;$$

(4) $\forall a \in R^{+}$,有负元素$a^{-1} \in R^{+}$,使

$$a \oplus a^{-1} = a \cdot a^{-1} = 1;$$

(5) $1 \otimes a = a^{1} = a;$

(6) $\lambda \otimes (\mu \otimes a) = \lambda \otimes a^{\mu} = (a^{\mu})^{\lambda} = a^{\lambda\mu} = (\lambda\mu) \otimes a;$

(7) $(\lambda + \mu) \otimes a = a^{\lambda+\mu} = a^{\lambda} a^{\mu} = a^{\lambda} \oplus a^{\mu}$

$$= (\lambda \otimes a) \oplus (\mu \otimes a);$$

(8) $\lambda \otimes (a \oplus b) = \lambda \otimes (ab) = (ab)^{\lambda} = a^{\lambda} b^{\lambda}$

$$= a^{\lambda} \oplus b^{\lambda} = (\lambda \otimes a) \oplus (\lambda \otimes b).$$

例1 数域 P 上的n维向量全体，按n维向量加法与n维向量的数乘构成数域 P 上的线性空间 P^n.

例2 数域 P 上 $m\times n$ 阶矩阵全体，按矩阵的加法和数乘构成 P 上的线性空间 $P^{m\times n}$.

例3 数域 P 上一元多项式全体（包括0）按照多项式的加法以及数与多项式的乘法构成 P 上的线性空间 $P[x]$.

例4 数域 P上次数不超过n的一元多项式（包括0）按照多项式的加法以及数与多项式的乘法也构成 P 上的线性空间 $P[x]_n$.

问： 数域 P上次数等于n的一元多项式（包括0）按照多项式的加法以及数与多项式的乘法也构成 P 上的线性空间吗？

例5 区间 $[a,b]$ 上全体连续实值函数全体按通常函数的加法和数与函数的乘法构成线性空间 $C[a,b]$.

例6 齐次线性方程组 $Ax=0$ 的所有解的集合构成数域 R 上的线性空间 $N(A)$，称为 $Ax=0$ 的**解空间**，或矩阵 A 的**核空间或零空间**，即

$$N(A)\equiv\{x\in R^n \mid Ax=0,\quad A\in R^{m\times n}\}$$

$$Ker(A)$$

问：$S\equiv\{x\in R^n \mid Ax=b,\ b\neq 0,\ A\in R^{m\times n}\}$?

例7 所有矩阵向量积 Ax 的集合构成数域 R 上的线性空间 $R(A)$，称为矩阵 A 的列空间或值域，也称为矩阵 A 的像， 即

$$R(A) \equiv \{y \in R^m \mid y = Ax, \quad x \in R^n, A \in R^{m\times n}\}$$

$$\mathrm{Im}(A)$$

例8 集合$V = \{x \mid x = [x_1, x_2, 1]^T，x_1, x_2 \in R\}$不是线性空间.

因为加法不封闭!

二、线性空间的基本性质

如果V是数域P上的**线性空间**，则

(1) 线性空间 V 中的**零向量** $\mathbf{0}$ 是唯一的.

(2) 线性空间 V 中的每个向量的**负向量**是唯一的.

(3) $0\cdot\alpha=\mathbf{0},\ k\cdot\mathbf{0}=\mathbf{0}.$

(4) 当 $k\alpha=\mathbf{0}$ 时，有 $k=0$ 或$\alpha=\mathbf{0}.$

(5) $(x-y)\alpha=x\alpha-y\alpha;\ x(\alpha-\beta)=x\alpha-x\beta.$

1.1.3 基、坐标、维数

一、基、坐标、维数的概念

定义1. 1. 6 给定线性空间 V，如果存在 V 中的一组向量 $\alpha_1, \cdots, \alpha_n$，满足：

（1）$\alpha_1, \cdots, \alpha_n$ 线性无关；

（2）V 中任意向量 α 都能由 $\alpha_1, \cdots, \alpha_n$ 线性表示.

即存在数 $x_1, \cdots, x_n \in P$，使

$$\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n$$

则向量组 $\alpha_1, \cdots, \alpha_n$ 就称为 V 的一组**基**, 系数 $x_1, \cdots, x_n$ 就称为向量 α 在此基下的**坐标**，基中的向量个数 n 称为线性空间 V 的**维数**，记为 $\dim V = n$.

例1 线性空间 $P[x]_n$ 是数域 P 上的n+1维空间，其基可取为 $1, x, x^2, \cdots, x^n$.

例2 线性空间 P^n 是数域 P 上的n维空间，其基可取为 $e_1, e_2, \cdots, e_n$.

例3 线性空间 $P^{m\times n}$ 是数域 P 上的m × n维空间，其基可取为 E_{ij}.

例4 向量空间 C 是实数域 R 上的二维空间，其基可取为 $\{1, i\}$，即

$$C = span\{1, i\} = \{a \cdot 1 + b \cdot i \mid a, b \in R\}.$$

同时向量空间 C 也是复数域 C 上的一维空间，其基可取为 $\{1\}$，即

$$C = span\{1\} = \{k \cdot 1 \mid k \in C\}.$$

说明:

（1）线性空间的基不一定存在. 例如：

$$V=\{0\},\ \dim\{0\}=0$$

$$V=P[x],\ \dim P[x]=\infty$$

（2）不存在有限个基向量的线性空间称为无限维线性空间. 本书只讨论有限维线性空间.

（3）n维线性空间任意n个线性无关的向量均构成一组基.

以下定理告诉我们，如何找到有限维线性空间的基.

定理（**基的扩张定理**）数域 P 上的 n 维**线性空间** V 中的任意一个线性无关向量组

$$\alpha_1, \alpha_2, \cdots, \alpha_r \quad (1 \le r \le n)$$

都可以扩充成 V 的一组基.

证明： 令 $W \equiv span\{\alpha_1,\alpha_2,\cdots,\alpha_r\}$,

则必有 $\alpha_{r+1} \in V$ 且 $\alpha_{r+1} \notin W$，使得

$$\alpha_1,\alpha_2,\cdots,\alpha_r,\alpha_{r+1} \text{ 线性无关.}$$

否则对任意 $\beta \in V - W$，都有 $\alpha_1,\alpha_2,\cdots,\alpha_r,\beta$ 线性相关. 这样 β 可由 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性表示，即 $\beta \in W$. 与 β 的取法矛盾.

重复上述过程，直至得到 V 的基 $\alpha_1,\alpha_2,\cdots,\alpha_n$.

定理1.1.1 数域 P 上的**线性空间** V 中的任意向量在给定基下的坐标是**唯一**的.

设 $\alpha_1,\cdots,\alpha_n$ 为 V 的一组基，$\forall\alpha\in V$，存在唯一的一组数 $x_1,\cdots,x_n\in P$，使

$$\alpha = x_1\alpha_1+\cdots+x_n\alpha_n$$

即 V 中的向量 α 与有序数组 $(x_1,\cdots,x_n)$ 一一对应.

这种一一对应保持了线性运算的对应关系，可以说线性空间 V 与 P^n 具有相同的结构，称为同构. 因为这种同构关系，**维数是有限维线性空间的唯一的本质特征**. 在同构的意义下，向量空间 P^n 并不只是线性空间 V 的一个特殊例子，而是所有的n维线性空间的代表. 只需主要研究元素为向量的线性空间 P^n.

由基的定义，在 n 维线性空间V中，任意n 个线性无关的向量都可以作为 V 的一组基.

对于不同的基，同一个向量的坐标是不同的.

那么，随着基的改变，同一个向量的坐标如何改变呢？

二、基变换和坐标变换

定义 设 $\alpha_1,\cdots,\alpha_n$和 $\beta_1,\cdots,\beta_n$ 是n 维线性空间V 的两个基，且存在可逆矩阵P，使得

$$[\beta_1,\cdots,\beta_n]=[\alpha_1,\cdots,\alpha_n]P.$$

则称上式为**基变换公式**，矩阵P 为基$\alpha_1,\cdots,\alpha_n$到基$\beta_1,\cdots,\beta_n$ 的**过渡矩阵**.

对于线性空间V，有

$$P=[\alpha_1,\cdots,\alpha_n]^{-1}[\beta_1,\cdots,\beta_n].$$

定理1.1.3 设n维线性空间V中元素α在基$\alpha_1,\cdots,\alpha_n$与基$\beta_1,\cdots,\beta_n$下的坐标分别为

$$x=[x_1,\cdots,x_n]^T,\ y=[y_1,\cdots,y_n]^T.$$

P为基$\alpha_1,\cdots,\alpha_n$到基$\beta_1,\cdots,\beta_n$的过渡矩阵，则成立**坐标变换公式**：

$$\boxed{x=Py}\ \text{即}\ \boxed{y=P^{-1}x.}$$

证明： 由于 $\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n$

$$= (\alpha_1, \cdots, \alpha_n)\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = (\alpha_1, \cdots, \alpha_n)x.$$

$$\alpha = \left[\beta_1, \cdots, \beta_n\right]y = (\alpha_1, \cdots, \alpha_n)Py.$$

因为坐标唯一，所以 $x = Py.$

由于 P 可逆，所以也有 $y = P^{-1}x.$

例5

已知$\alpha_1,\alpha_2,\alpha_3$ 线性空间V^3的一组基，向量组β_1,β_2,β_3 满足

$$\beta_1+\beta_3=\alpha_1+\alpha_2+\alpha_3,\beta_1+\beta_2=\alpha_2+\alpha_3,\beta_2+\beta_3=\alpha_1+\alpha_3.$$

（1）证明：向量组β_1,β_2,β_3也是一组基.

（2）求由基β_1,β_2,β_3 到基$\alpha_1,\alpha_2,\alpha_3$ 的过渡矩阵.

（3）求向量$\alpha=\alpha_1+2\alpha_2-\alpha_3$ 在基β_1,β_2,β_3 下的坐标.

解：（1）由已知条件，有

$$(\beta_1, \beta_2, \beta_3)\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} = (\alpha_1, \alpha_2, \alpha_3)\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

由于上式两边矩阵均可逆，所以$\beta_1, \beta_2, \beta_3$线性无关，
即$\beta_1, \beta_2, \beta_3$也是一组基.

(2)

$$(\alpha_1,\alpha_2,\alpha_3)=(\beta_1,\beta_2,\beta_3)\begin{pmatrix}1&1&0\\0&1&1\\1&0&1\end{pmatrix}\begin{pmatrix}1&0&1\\1&1&0\\1&1&1\end{pmatrix}^{-1}$$

$$=(\beta_1,\beta_2,\beta_3)\begin{pmatrix}0&1&0\\-1&-1&2\\1&0&0\end{pmatrix}.$$

所以，由基β_1,β_2,β_3 到基$\alpha_1,\alpha_2,\alpha_3$ 的过渡矩阵为

$$\begin{pmatrix}0&1&0\\-1&-1&2\\1&0&0\end{pmatrix}.$$

（3）

$$\alpha=\alpha_1+2\alpha_2-\alpha_3=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix}1\\2\\-1\end{pmatrix}$$

$$=(\beta_1,\beta_2,\beta_3)\begin{pmatrix}0&1&0\\-1&-1&2\\1&0&0\end{pmatrix}\begin{pmatrix}1\\2\\-1\end{pmatrix}$$

$$=(\beta_1,\beta_2,\beta_3)\begin{pmatrix}2\\-5\\1\end{pmatrix}.$$

所以，向量$\alpha=\alpha_1+2\alpha_2-\alpha_3$ 在基β_1,β_2,β_3 下的坐标为$(2,-5,1)$.

整体有时太庞大，所以我们经常希望能够“通过部分来获知整体”，从而达到“解剖麻雀”的效果. 对线性空间的研究亦是如此. 我们希望通过对线性子空间的研究，能够更加深刻地揭示整个线性空间的结构.

主要内容

- 子空间的概念
- 子空间的判定
- 子空间的构造
- 子空间的运算
- 维数定理

1.1.4 子空间的定义

定义1.1.8 设 U是线性空间 V的非空子集. 如果 U在 V 中规定的加法和数乘运算下构成线性空间，则称 U是 V的（线性）子空间.

例1 {0}和 V为V的平凡子空间.

例2 $P[x]_n$ 为$P[x]$的子空间.

判定非空集合是否为线性空间，要验算加法和数乘运算的封闭性，以及8条运算律. 至于判定线性空间的子集是否为线性空间，可用下面定理判定.

定理1.1.4（子空间判别法）数域 P 上的**线性空间** V 的非空子集 U 是 V 的子空间的充要条件是 U 对 V 中的两种运算都封闭，即

（i）　对任意的 $\alpha,\beta\in U$，有 $\alpha+\beta\in \mathbf{U}$.

（ii）对任意的 $x\in \mathbf{P}$，$\alpha\in \mathbf{U}$，有 $x\alpha\in \mathbf{U}$.

例3 设$A \in M^{m\times n}(R)$, $Ax=0$的解空间$N(A)$($\ker A$)是R^n的线性子空间. A的值域$R(A)$($\mathrm{Im}\, A$)是R^m的线性子空间.

例4 R^3中过原点的直线是R^3的一个子空间.

例5 集合$T_1=\{x \mid x=(x_1,x_2,0)^T, x_1,x_2\in R\}$是向量空间. 它是$R^3$在$\mathrm{O}x_1x_2$平面上的**投影子空间**.

例6 已知 V 是数域 P 上的线性空间，α、$\beta\in V$，则集合

$$W=\left\{x=\lambda\alpha+\mu\beta \;\middle|\; \lambda,\mu\in P\right\}$$

是 V 的一个子空间. 称为**由向量** α、β **所生成的子空间**，记为 $span(\alpha,\beta)$ 或 $L(\alpha,\beta)$.

定理1.1.5 由线性空间 V 中的**向量组** $\alpha_1,\cdots,\alpha_s$ **所生成的线性空间**

$$W=\left\{x=\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_s\alpha_s \;\middle|\; \lambda_1,\lambda_2,\cdots,\lambda_s\in R\right\}$$

记作 $span(\alpha_1,\alpha_2,\cdots,\alpha_s)$ 或 $L(\alpha_1,\alpha_2,\cdots,\alpha_s)$.

1.1.5 子空间的交与和

定理1.1.6 设 V_1,V_2 是数域 P 上线性空间 V 的两个子空间，则它们的**交** $V_1\cap V_2$ 也是 V 的子空间.

定理1.1.7 设 V_1,V_2 是数域 P 上线性空间 V 的两个子空间，则集合（称为 V_1 与 V_2 的**和**）

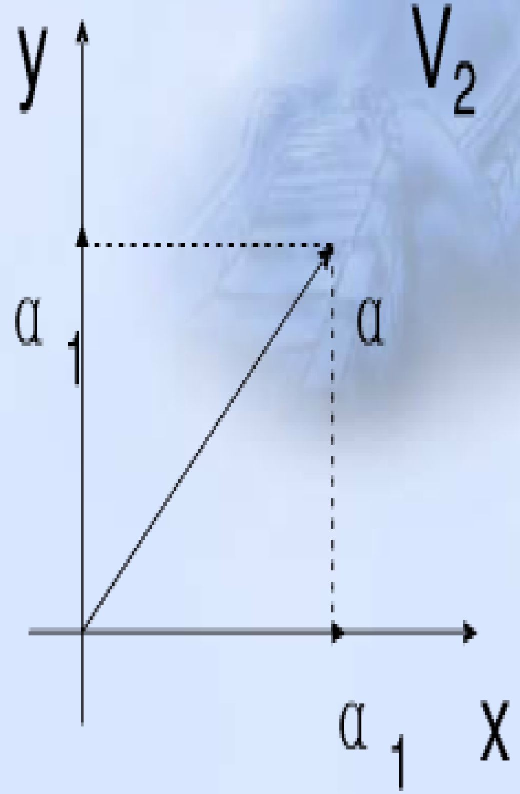
$$V_1+V_2\equiv\{\alpha_1+\alpha_2\mid\alpha_1\in V_1,\alpha_2\in V_2\}$$

也是 V 的子空间.

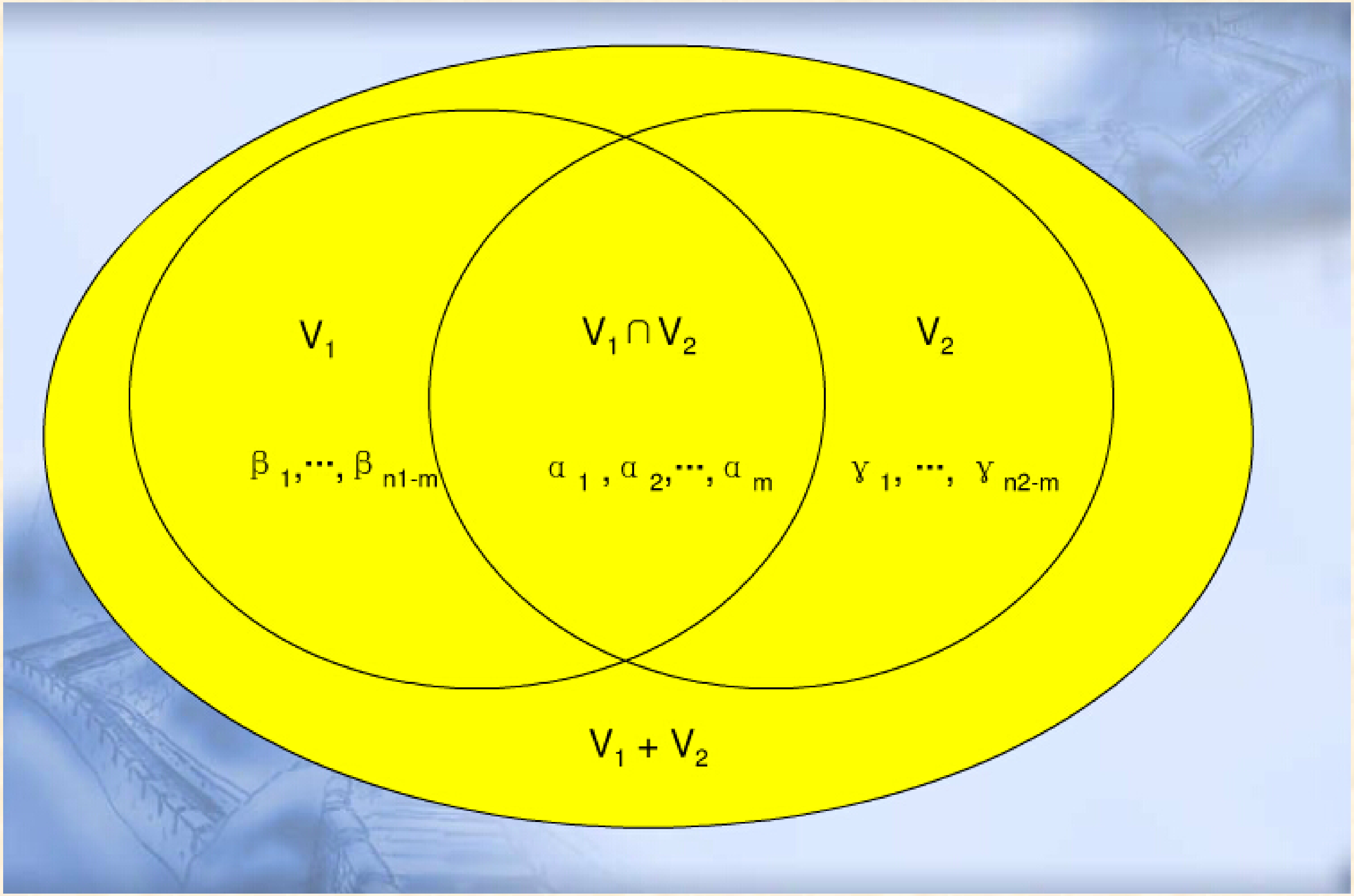
为什么要研究子空间的**和**？

为什么不研究子空间的**并**？

例：二维平面V_2中，W=x轴，V=y轴均为V的子空间. 如下图所示，向量$\alpha_1 \in W$，$\alpha_2 \in V$，但$\alpha_1+\alpha_2$却不在$W \cup V$中.



矩阵论(MATRIX THEORY)



可以验证，子空间的交与和具有下列运算律：

（i）交换律 $$V_1 \cap V_2 = V_2 \cap V_1$$

$$V_1 + V_2 = V_2 + V_1$$

（ii）结合律 $$(V_1 \cap V_2) \cap V_3 = V_1 \cap (V_2 \cap V_3)$$

$$(V_1 + V_2) + V_3 = V_1 + (V_2 + V_3)$$

根据归纳法可知，$\bigcap\limits_{i=1}^{m} V_i$ 和 $\sum\limits_{i=1}^{n} V_i$ 都是 V 的子空间.

例1 设 V_1,V_2是线性空间 V 的子空间，且

$$V_1 = span(\alpha_1,\cdots,\alpha_s),$$

$$V_2 = span(\beta_1,\cdots,\beta_t),$$

则

$$V_1 + V_2 = span(\alpha_1,\cdots,\alpha_s,\beta_1,\cdots,\beta_t).$$

定理1.1.8 **（维数公式）** 设 V_1, V_2 是数域 P 上线性空间 V 的两个有限维子空间，则它们的**交**与**和**都是有限维的，并且

$$\dim(V_1+V_2)=\dim(V_1)+\dim(V_2)-\dim(V_1\cap V_2).$$

证明：

令 $\dim(V_1)=s, \dim(V_2)=t, \dim(V_1\cap V_2)=r.$

取 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 为 $V_1\cap V_2$ 的基，并分别扩充成 V_1, V_2 的基，即

$$V_1=span(\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_{r+1},\cdots,\beta_s),$$

$$V_2=span(\alpha_1,\alpha_2,\cdots,\alpha_r,\gamma_{r+1},\cdots,\gamma_t),$$

所以

$$V_1+V_2=span(\alpha_1,\cdots,\alpha_r,\beta_{r+1},\cdots,\beta_s,\gamma_{r+1},\cdots,\gamma_t).$$

设 $k_1\alpha_1+\cdots+k_r\alpha_r+p_{r+1}\beta_{r+1}+\cdots+p_s\beta_s+q_{r+1}\gamma_{r+1}+\cdots+q_t\gamma_t=0$

则有

$$\alpha=k_1\alpha_1+\cdots+k_r\alpha_r+p_{r+1}\beta_{r+1}+\cdots+p_s\beta_s\in V_1$$

又

$$\alpha=-(q_{r+1}\gamma_{r+1}+\cdots+q_t\gamma_t)\in V_2$$

所以 $\alpha\in V_1\cap V_2$.

从而可令 $\alpha=l_1\alpha_1+\cdots+l_r\alpha_r$,

则 $l_1\alpha_1+\cdots+l_r\alpha_r=-(q_{r+1}\gamma_{r+1}+\cdots+q_t\gamma_t)$.

由于 $\alpha_1,\cdots,\alpha_r,\gamma_{r+1},\cdots,\gamma_t$ 线性无关，所以

$$l_1=\cdots=l_r=q_{r+1}=\cdots=q_t=0.$$

因此 $\alpha=0.$

即 $\alpha=k_1\alpha_1+\cdots+k_r\alpha_r+p_{r+1}\beta_{r+1}+\cdots+p_s\beta_s=0.$

再由 $\alpha_1,\cdots,\alpha_r,\beta_{r+1},\cdots,\beta_s$ 线性无关，所以

$$k_1=\cdots=k_r=p_{r+1}=\cdots=p_s=0.$$

从而 $\alpha_1,\cdots,\alpha_r,\beta_{r+1},\cdots,\beta_s,\gamma_{r+1},\cdots,\gamma_t$ 线性无关，所以是 V_1+V_2 的一组基．即证维数公式成立.

由维数公式可知，子空间和的维数要比维数的和小.例如三维空间V_3中，两张通过原点的不同平面之和是整个三维空间V_3，而其维数之和是4，由此说明这两张平面的交是一维的直线.

例2 设 $\alpha_1=(2,1,3,1)^T,\alpha_2=(-1,1,-3,1)^T,$

$\beta_1=(4,5,3,-1)^T,\beta_2=(1,5,-3,1)^T,$

$V_1=span(\alpha_1,\alpha_2),V_2=span(\beta_1,\beta_2).$

求 $V_1\cap V_2$、V_1+V_2 的基与维数.

解： 由题意知

$$\dim(V_1)=2,\ \dim(V_2)=2$$

由例1 知 $V_1+V_2=span(\alpha_1,\alpha_2,\beta_1,\beta_2)$.

由

$$\begin{pmatrix} 2 & -1 & 4 & 1 \\ 1 & 1 & 5 & 5 \\ 3 & -3 & 3 & -3 \\ 1 & 1 & -1 & 1 \end{pmatrix} \xrightarrow{ERT} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & \frac{5}{3} \\ 0 & 0 & 1 & \frac{2}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$ 行最简形

知 $\alpha_1,\alpha_2,\beta_1$ 线性无关，从而是 V_1+V_2 的一组基

$\dim(V_1+V_2)=3.$

由维数定理知，$\dim(V_1 \cap V_2) = 1.$

设 $\alpha \in V_1 \cap V_2$ ，则 $\alpha \in V_1$， $\alpha \in V_2$

所以可令 $k_1\alpha_1 + k_2\alpha_2 = \alpha = l_1\beta_1 + l_2\beta_2$

即
$$\begin{cases} 2k_1 - k_2 - 4l_1 - l_2 = 0 \\ k_1 + k_2 - 5l_1 - 5l_2 = 0 \\ 3k_1 - 3k_2 - 3l_1 + 3l_2 = 0 \\ k_1 + k_2 + l_1 - l_2 = 0 \end{cases}$$

解关于 k_1， k_2， l_1, l_2 的齐次方程.

由

$$\begin{pmatrix} 2 & -1 & -4 & -1 \\ 1 & 1 & -5 & -5 \\ 3 & -3 & -3 & 3 \\ 1 & 1 & 1 & -1 \end{pmatrix} \xrightarrow{ERT} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -\frac{5}{3} \\ 0 & 0 & 1 & \frac{2}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

行最简形

得 $k_1 = 0,\ k_2 = \frac{5}{3} l_2,\ l_1 = -\frac{2}{3} l_2,$

$$\boxed{\beta_2 = -\frac{5}{3}\alpha_2 + \frac{2}{3}\beta_1}$$

因此 $\alpha = k_1\alpha_1 + k_2\alpha_2 = \frac{5}{3} l_2 \cdot \alpha_2.$

所以 $V_1 \cap V_2$ 的基为 $\alpha_2.$

练习：已知 $\alpha_1=(-1,3,2,1)^T,\alpha_2=(1,0,0,2)^T$，

$$\beta_1=(0,1,1,3)^T,\beta_2=(-3,2,1,-6)^T$$

且 $V_1=span\{\alpha_1,\alpha_2\},V_2=span\{\beta_1,\beta_2\}$，

求 V_1+V_2 与 $V_1\cap V_2$ 的基和维数.

解：因为

$$V_1+V_2=span\{\alpha_1,\alpha_2\}+span\{\beta_1,\beta_2\}=span\{\alpha_1,\alpha_2,\beta_1,\beta_2\},$$

而

$$(\alpha_1,\alpha_2,\beta_1,\beta_2)=\begin{pmatrix}-1&1&0&-3\\3&0&1&2\\2&0&1&1\\1&2&3&-6\end{pmatrix}\xrightarrow{\text{初等行变换}}\begin{pmatrix}1&0&0&1\\0&1&0&-2\\0&0&1&-1\\0&0&0&0\end{pmatrix}$$

由于 $\alpha_1,\alpha_2,\beta_1$ 是向量组 $\alpha_1,\alpha_2,\beta_1,\beta_2$ 的一个极大线性无关组，则和空间的维数是3，基为 $\alpha_1,\alpha_2,\beta_1$

且 $\beta_2=\alpha_1-2\alpha_2-\beta_1$.

由行最简形知 $\dim(V_1)=2, \dim(V_2)=2,$

又

$$\dim(V_1+V_2)=\dim V_1+\dim V_2-\dim(V_1\cap V_2).$$

故 $\dim(V_1\cap V_2)=1$。 由 $\beta_2=\alpha_1-2\alpha_2-\beta_1$ 得

$$\xi=\alpha_1-2\alpha_2=\beta_1+\beta_2=(-3,3,2,-3)^T\in V_1\cap V_2.$$

所以 $(-3,3,2,-3)^T$ 为 $V_1\cap V_2$ 的一组基.

由子空间 V_1+V_2 的定义可知，若 $x_1\in V_1, y\in V_2$，则 V_1+V_2 中任一向量z可以表示为 $z=x+y$. 但是，一般来说，这种表示并不是唯一的. 例如：在 R^3 中

若$V_1=span\{(1,0,0),(0,1,0)\}$， $V_2=span\{(1,1,0)\}$

则有$z=(1,2,0)\in V_1+V_2$, 但

$$(1,2,0)=(1,2,0)+(0,0,0)\in V_1+V_2$$

$$(1,2,0)=(0,1,0)+(1,1,0)\in V_1+V_2$$

这说明和空间的向量z的表示法不唯一！

$$\dim(V_1+V_2)=\dim(V_1)+\dim(V_2)-\dim(V_1\cap V_2).$$

由维数公式：和空间的维数不大于子空间维数之和.

定义1.1.10 设 V_1, V_2是数域 P上的线性空间V 的两个子空间，如果

$$V_1\cap V_2=\{0\},$$

则称 V_1+V_2为 V_1 与 V_2 的**直和**，记作$V_1\oplus V_2$.

显然直和的概念可以推广到多个子空间的情形.

定理1.1.9 设 V_1, V_2 是数域 P 上的线性空间 V 的两个子空间，则下列命题是等价的：

（1） V_1+V_2 是直和；

（2） $\dim(V_1+V_2)=\dim(V_1)+\dim(V_2)$；

（3） 和 V_1+V_2 中零向量的表示法唯一，即若

$$\alpha_1+\alpha_2=0 \quad (\alpha_1\in V_1,\ \alpha_2\in V_2),$$

则 $\alpha_1=\alpha_2=0.$

（4） 和 V_1+V_2 中每个向量的表示法是唯一的.

证明： $(1)\Rightarrow(2)$. 根据维数公式显然成立.

$(2)\Rightarrow(1)$. 根据维数公式 $\dim(V_1\cap V_2)=0$.

所以 $V_1\cap V_2=\{0\}$.

$(1)\Rightarrow(4)$. 设存在向量 $\alpha\in V_1+V_2$，有

$$\alpha=\alpha_1+\alpha_2=\alpha_1'+\alpha_2',\ \alpha_1,\ \alpha_1'\in V_1,\alpha_2,\ \alpha_2'\in V_2$$

由于 $\alpha_1-\alpha_1'\in V_1,\alpha_2'-\alpha_2\in V_2$,且$\alpha_1-\alpha_1'=\alpha_2'-\alpha_2$,

所以$\alpha_1-\alpha_1'=\alpha_2'-\alpha_2\in V_1\cap V_2=\{0\}$.

从而$\alpha_1-\alpha_1'=\alpha_2'-\alpha_2=0$.

(4) $\Rightarrow$ (3). 显然成立.

(3) $\Rightarrow$ (1). 对任意向量 $\alpha \in V_1 \cap V_2$，有

$$0 = \alpha + (-\alpha), \alpha \in V_1, \ -\alpha \in V_2.$$

根据（3），零向量的表示是唯一的，因此

$$\alpha = -\alpha = 0.$$

例3 设S, K 分别是 n 阶实对称矩阵和反对称矩阵的全体. 显然容易证明 S, K 均为线性空间 $R^{n\times n}$ 的子空间. 试证明 $R^{n\times n} = S \oplus K$.

证明：因为任意实方阵可以分解为一个实对称矩阵和一个实反对称矩阵的和，即

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T), \quad A \in R^{n\times n}$$

又 $\dim(R^{n\times n}) = n^2 = n(n+1)/2 + n(n-1)/2$

$= \dim(S) + \dim(K)$，根据定理1.1.9可知结论成立.

定理 （**直和分解**）设 V_1 是数域 P 上的线性空间 V 的一个子空间，则一定存在 V 的另一个子间 V_2，使得线性空间 V 具有**直和分解**

$$V = V_1 \oplus V_2$$

则称 V_1 和 V_2 是一对**互补的子空间**，或者 V_1 是 V_2 的**补子空间**.

显然直和分解可以推广到多个子空间的情形.

例如： R^3 的直和分解与非直和分解.

设 $V_1 = \{(x_1, x_2, x_3) \mid x_3 = 0, x_1, x_2 \in R\}$,

$V_2 = \{(x_1, x_2, x_3) \mid x_1 = 0, x_2 = 0, x_3 \in R\}$,

$V_3 = \{(x_1, x_2, x_3) \mid x_1 = 0, x_2, x_3 \in R\}$.

则 $R^3 = V_1 + V_2$ 是直和分解,

$R^3 = V_1 + V_3$ 不是直和分解.

注意：**子空间的补子空间未必是唯一的**，也就是说**线性空间的直和分解未必是唯一的.**

例如：若 $\alpha_1=(1,0,0)^T$， $\alpha_2=(0,1,0)^T$，

$$\beta_1=(0,0,1)^T,\quad \beta_2=(0,1,1)^T.$$

显然，$U=span(\alpha_1,\alpha_2)$ 是 R^3 的 一个子空间，

且 $span(\beta_1)$ 和 $span(\beta_2)$ 都 是 U 的补子空间.

1.2 线性变换

线性变换是线性空间的核心内容，反映的是线性空间中元素间的一种基本联系，体现出一种“动态的”或者“直观的”视角.

借助基的概念，可在线性变换与矩阵之间建立一一对应关系，因此通俗地讲“变换即矩阵”. 这同时也意味着线性变换的运算可以转化为矩阵的运算.

1.2.1 线性变换的定义

例1（**旋转变换**或**Givens变换**）将线性空间R^2中的所有向量均绕原点**顺时针**旋转角θ，则**像**(η_1,η_2)与**原像**(ξ_1,ξ_2)之间的关系为

$$\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}.$$

例2（**反射变换**或**Householder变换**）向量l_θ为R^2中与x轴正向夹角为θ且过原点的向量. 将R^2中的所有向量均以向量l_θ为轴做反射，则**像**(η_1,η_2)与**原像**(ξ_1,ξ_2)之间的关系为

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

矩阵论(MATRIX THEORY)

从几何上看，图形经过旋转变换或反射变换后只是位置改变了，形状和大小都没有改变，所有的长度、角度都保持不变，直线仍然变成了直线，三角形、长方形、正方形、平行四边形、圆仍然变成了三角形、长方形、正方形、平行四边形、圆，也就是说变换前后的图形是全等的.

从映射的角度看，旋转变换或反射变换对应的显然是**同构映射**.

例3（**伸缩变换**）将 R^2 中的单位圆沿坐标轴方向分别拉伸 a 倍和 b 倍，所得图形为椭圆，则**像**(η_1,η_2)与**原像**(ξ_1,ξ_2)之间的关系为

$$\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}=\begin{pmatrix}a & 0\\ 0 & b\end{pmatrix}\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix}.$$

从几何上看，图形经过**伸缩变换**后位置没有改变，角度也没有改变，但形状和大小都略有变化，所有的长度也改变了. 虽然直线仍然变成了直线，三角形、长方形、平行四边形仍然变成了三角形、长方形、平行四边形，但圆变成了椭圆、正方形变成了长方形. 也就是说变换前后的图形未必是全等的.

从映射的角度看，**伸缩变换**对应的显然不再是**同构映射**. 但是**线性关系经过变换仍然得到保留**.

定义1.2.1 设 V 是数域 P 上的线性空间，映射（未必是双射）$T:V\to V$ 称为 V 上的**线性变换**或**线性算子**，如果对 V 中的任意两个向量 α、$\beta\in V$ 和任意数 $k\in P$，都有

（i）（可加性） $T(\alpha+\beta)=T(\alpha)+T(\beta)$；

（ii）（齐次性） $T(k\alpha)=kT(\alpha)$.

则称 $T(\alpha)$ 为 α 在 T 下的**像**，而 α 是 $T(\alpha)$ 的**原像**.

例4 由下式

$$T(x) = Ax, \quad x \in R^n, A \in R^{n\times n}$$

确定的映射 $T: R^n \to R^n$ 是线性变换.

当矩阵 A 是对角矩阵时，该变换是一种伸缩变换.
特别地，如果 A 是单位矩阵，该变换即为恒等变换.

例5（**标量变换**）由下式

$$T(x)=\lambda x,\quad x\in V,\lambda\in P$$

确定的线性空间 V 到其自身的映射 $T:V\to V$ 是线性变换. 这里 λ 称为**线性变换 T 的特征值**, 非零向量 $x\neq 0$ 称为 T **的对应于特征值 λ 的特征向量**.

例6 数域 $\boldsymbol{R}$ 上的所有无限次可导实函数的集合 $\boldsymbol{V}$ 是一个线性空间，则由下式

$$\boldsymbol{D(f) = f'}, \quad \boldsymbol{f \in V}$$

确定的微商变换 $\boldsymbol{D : V \to V}$ 是 $\boldsymbol{V}$ 上的一个线性变换.

例7 闭区间 $[a,b]$ 上的所有实连续函数的集合 $C[a,b]$ 构成 $\boldsymbol{R}$ 上的一个线性空间，则由下式

$$J(f)=\int_a^t f(u)\mathrm{d}u,\quad f\in C[a,b]$$

确定的积分变换 $J: C[a,b]\to C[a,b]$ 是 $C[a,b]$ 上的一个线性变换.

例 6和例 7 表明，微积分的两个基本运算（微分和积分），从变换的角度看都是线性变换（或线性算子），由此可知线性变换在理论与应用中有着广泛的应用.

1.2.2 线性变换的性质

定理 如果 $T:V\to V$是线性变换，则

(1) $T(0)=0$; 零向量对应零向量

(2) $T(-\alpha)=-T(\alpha)$; 负向量对应负向量

(3) $T\left(\sum_{i=1}^{m}\lambda_i\alpha_i\right)=\sum_{i=1}^{m}\lambda_i T(\alpha_i)$. 叠加原理

(4) 原像线性相关 $\Rightarrow$ 像线性相关.

注：线性变换可能将线性无关的向量组变成线性相关的向量组，如 零变换.

1.2.3 线性变换的运算

定理 如果 $L(V)$ 表示 V 上的所有线性变换的集合，并且对任意 T、T_1、$T_2 \in L(V)$, $\alpha \in V, k \in P$

(1) $(T_1 + T_2)(\alpha) \equiv T_1(\alpha) + T_2(\alpha)$;

(2) $(kT)(\alpha) \equiv kT(\alpha)$;

(3) $(T_1T_2)(\alpha) \equiv (T_1(T_2(\alpha))$;

则可以验证，$T_1 + T_2, kT, T_1T_2$ 都是线性变换，因此 $L(V)$ 也是数域 P 上的线性空间.

1.2.4 线性变换的值域与核

设 T 是数域 P 上的线性空间 V 上的线性变换. 令

$$\mathrm{Im}(T) \equiv R(T) \equiv \{ T(\alpha) \mid \forall \alpha \in V \},$$

$$Ker(T) \equiv N(T) \equiv \{ \alpha \in V \mid T(\alpha) = 0 \},$$

易知 $\mathrm{Im}(T)$ 和 $Ker(T)$ 都是 V 的子空间.

称 $\mathrm{Im}(T)$ 是**线性变换 T 的值域**，而 $Ker(T)$ 是**线性变换的核**. $\mathrm{Im}(T)$ 的维数称为 T 的**秩**，$Ker(T)$ 的维数称为 T 的**零度**.

设 T 是数域 P 上的线性空间 V 上的线性变换.

令 $\alpha_1,\alpha_1,\cdots,\alpha_n$ 为 V 的一组基，则

$$\mathrm{Im}(T)=span(T(\alpha_1),T(\alpha_2),\cdots,T(\alpha_n)).$$

例1 已知$R^{2\times 2}$ 的线性变换, $T(X)=MX-XM$, $M=\begin{pmatrix}1 & 2\\ 0 & 3\end{pmatrix}$

求：$R(T)$ 与$N(T)$ 的基与维数.

解：取$R^{2\times 2}$ 的自然基$E_{11}, E_{12}, E_{21}, E_{22}$, 则

$$T(E_{11})=\begin{bmatrix}0 & -2\\ 0 & 0\end{bmatrix},\quad T(E_{12})=\begin{bmatrix}0 & -2\\ 0 & 0\end{bmatrix},$$

$$T(E_{21})=\begin{bmatrix}2 & 0\\ 2 & -2\end{bmatrix},\quad T(E_{22})=\begin{bmatrix}0 & 2\\ 0 & 0\end{bmatrix}.$$

显然, $T(E_{11}), T(E_{21})$ 是一个最大线性无关组.所以,
$dim(R(T))=2$, 且$T(E_{11}), T(E_{21})$ 是$R(T)$ 的一组基.

设$X=\begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}\in N(T)$, 由

$$T(X)=MX-XM=\begin{bmatrix} 2x_3 & -2x_1-2x_2+2x_4 \\ 2x_3 & -2x_3 \end{bmatrix}=0,$$

解得$X=\left\{\begin{bmatrix} -c_1+c_2 & c_1 \\ 0 & c_2 \end{bmatrix}\middle| \quad c_1,c_2\in R\right\}.$

所以, $dim(N(T))=2$, 且$\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

是$N(T)$ 的一组基.

例2 $P[x]_{n-1}$ 为变量 x 的次数不超过 $n-1$ 的实系数多项式构成的空间（n 维），并设求导变换为 D，则

$$D\left(P[x]_{n-1}\right)=\left(P[x]_{n-1}\right)'=P[x]_{n-2}$$

故 $\mathrm{Im}(D)=P[x]_{n-2}$

$$Ker(D)=P$$

且

$$\dim\left(Ker(D)\right)+\dim\left(\mathrm{Im}(D)\right)=n=\dim\left(P[x]_{n-1}\right).$$

设 T 是数域 P 上的 n 维线性空间 V 上的线性变换.

注意:

虽然 $\dim \mathrm{Im}(T) + \dim Ker(T) = n$，但一般有

$$\mathrm{Im}(T) + Ker(T) \neq V.$$

1.2.5 线性变换的矩阵表示

定义 n 维线性空间 V 上的线性变换

$$T: V \to V$$

将 V 的基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 映射为 $T(\alpha_1), T(\alpha_2), \cdots, T(\alpha_n)$.

由于 $T(\alpha_i)$ 仍然是基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的线性组合，令

$$T(\alpha_i) = a_{1i}\alpha_1 + a_{2i}\alpha_2 + \cdots + a_{ni}\alpha_n \quad (i = 1, 2, \cdots, n)$$

因此

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

$$=(\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$\equiv(\alpha_1,\alpha_2,\cdots,\alpha_n)A.$$

这里，矩阵 A 称为**线性变换** T 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下**的矩阵表示**.

因此线性变换与方阵之间可以建立一一对应关系.

定理1.2.6

设A, B 分别是线性变换T_1 和T_2 在基$\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵, 那么在此基下有:

（1）$T_1 + T_2$ 在基$\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵为$A + B$;

（2）kT_1 在基$\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵为kA;

（3）T_1T_2 在基$\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵为AB;

（4）若T_1 可逆, 则T_1^{-1} 在基$\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵为A^{-1}.

注: 定理中的(1)(2)说明了$L(V)$与$P^{n\times n}$同构.

定理1.2.4 设线性变换 T 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵为为 A，且

向量 α 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的坐标为 $x=\left(x_1,x_2,\cdots,x_n\right)^T$

$T(\alpha)$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的坐标为 $y=\left(y_1,y_2,\cdots,y_n\right)^T$

则

$$\boxed{y = A\,x.}$$

对 V 中的任意向量 $\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n$

显然其在线性变换 T 下的像为

$$T(\alpha) = T(x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n)$$

$$= x_1T(\alpha_1) + x_2T(\alpha_2) + \cdots + x_nT(\alpha_n)$$

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

$$= (\alpha_1, \alpha_2, \cdots, \alpha_n)Ax \equiv (\alpha_1, \alpha_2, \cdots, \alpha_n)y.$$

因此原像与像（在给定基下）的坐标变换公式为

$$\boxed{y = A\,x.}$$

例1 R^3 中的投影变换

$$P:(x,y,z)\to(x,y,0),$$

在基 i, j, k 下的矩阵为

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

例 2 $P[x]_n$ 中的微商变换

$$D(f)=f',\quad f\in P[x]_n$$

在基 $1, x, \cdots, x^n$ 下的矩阵为

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

例3

已知$P_2[t]$上线性变换

$$T(a_0+a_1t+a_2t^2)=(a_0-a_1)+(a_1-a_2)t+(a_2-a_0)t^2.$$

求T 在基$1,t,t^2$ 下的矩阵.

解 因为

$$T(1)=1-t^2,\quad T(t)=-1+t,\quad T(t^2)=-t+t^2,$$

所以T 在基$1,t,t^2$ 下的矩阵为 $A=\begin{bmatrix}1&-1&0\\0&1&-1\\-1&0&1\end{bmatrix}.$

例4 在 R^3 中 T 将基 $\alpha_1,\alpha_2,\alpha_3$ 变为基 β_1,β_2,β_3, 其中

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

$$\beta_1=(1,-1,0)^T,\quad \beta_2=(0,1,-1)^T,\quad \beta_3=(0,3,-2)^T$$

(1) 求 T 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵;

(2) 求向量 $\xi=(1,2,3)^T$ 及 $T(\xi)$ 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的坐标.

解： (1)设 T 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵为 A，则

$$(\beta_1,\beta_2,\beta_3)=T(\alpha_1,\alpha_2,\alpha_3)=(\alpha_1,\alpha_2,\alpha_3)A$$

$$\Rightarrow A=(\alpha_1,\alpha_2,\alpha_3)^{-1}(\beta_1,\beta_2,\beta_3)$$

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

$$=\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}.$$

设$\xi = x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3 = (\alpha_1, \alpha_2, \alpha_3)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

则$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (\alpha_1, \alpha_2, \alpha_3)^{-1}\xi$

$$= \begin{pmatrix} 2 & 1 & 2 \\ -1 & 0 & -1 \\ -1 & -1 & -2 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 10 \\ -4 \\ -9 \end{pmatrix}.$$

设$T(\xi)$在基$\alpha_1,\alpha_2,\alpha_3$下的坐标为$(y_1,y_2,y_3)^T$.

则 $$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = A\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$=\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 10 \\ -4 \\ -9 \end{pmatrix}=\begin{pmatrix} 23 \\ -32 \\ -13 \end{pmatrix}.$$

当基改变后，线性变换的表示矩阵是否改变呢？

设 T 为 n 维向量空间 V 上的线性变换，对于 V 的基 $x_1,x_2,\cdots,x_n$ 和 $y_1,y_2,\cdots,y_n$ 的矩阵表示分别是 A 和 B，并且基 $x_1,x_2,\cdots,x_n$ 到基 $y_1,y_2,\cdots,y_n$ 的过渡矩阵为 P，即有

$$(y_1,y_2,\cdots,y_n)=(x_1,x_2,\cdots,x_n)P.$$

$$(x_1,x_2,\cdots,x_n)=(y_1,y_2,\cdots,y_n)P^{-1}.$$

则 $$(y_1, y_2, \cdots, y_n)B = T(y_1, y_2, \cdots, y_n)$$

$$= T(x_1, x_2, \cdots, x_n)P$$

$$= (Tx_1, Tx_2, \cdots, Tx_n)P$$

$$= (x_1, x_2, \cdots, x_n)AP$$

$$= (y_1, y_2, \cdots, y_n)P^{-1}AP,$$

故 $B = P^{-1}AP$，或

$$A = PBP^{-1}.$$

定理1.2.5 设 T 为 n 维向量空间 V 上的线性变换，对于 V 的基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 和 $\beta_1,\beta_2,\cdots,\beta_n$ 的矩阵表示分别是 A 和 B，且基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 到基 $\beta_1,\beta_2,\cdots,\beta_n$ 的过渡矩阵为 P，则矩阵 A 与 B 相似，即存在可逆矩阵 P 使得

$$B = P^{-1}AP.$$

据此定理可知，同一个线性变换下在不同基下矩阵表示是相似的. 反之，若两矩阵相似，则它们可看成同一线性变化在不同基下的矩阵.

例5 设 T 是 R^3 中的线性变换

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

$$\beta_1=(1,1,1)^T,\ \beta_2=(1,1,0)^T,\ \beta_3=(1,0,0)^T$$

是 R^3 的两组基，且

$$T(\beta_1)=(1,0,0)^T, T(\beta_2)=(2,1,0)^T, T(\beta_3)=(3,2,1)^T$$

求 T 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵.

解：$(\beta_1,\beta_2,\beta_3)=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix}1&0&0\\1&1&0\\1&1&1\end{pmatrix},$

$\Rightarrow \alpha_1,\alpha_2,\alpha_3$到$\beta_1,\beta_2,\beta_3$的过渡矩阵为

$$P=\begin{pmatrix}1&0&0\\1&1&0\\1&1&1\end{pmatrix}.$$

$$\text{而}\,T(\beta_1,\beta_2,\beta_3)=(\beta_1,\beta_2,\beta_3)\begin{pmatrix}0&0&1\\0&1&1\\1&1&1\end{pmatrix},$$

故T在β_1,β_2,β_3下的矩阵为

$$B=\begin{pmatrix}0&0&1\\0&1&1\\1&1&1\end{pmatrix}.$$

因此$\boldsymbol{T}$在$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$下的矩阵为

$$\boldsymbol{A}=\boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1}=\begin{pmatrix} 0 & -1 & 1 \\ -1 & -1 & 2 \\ -1 & -1 & 3 \end{pmatrix}.$$

例6* 设$B=\begin{pmatrix} 0 & 1 \\ 4 & 0 \end{pmatrix}\in R^{2\times 2}$, 在$R^{2\times 2}$中定义线性变换为

$$T(X)=XB$$

试求$R^{2\times 2}$的一组基是T在所求基下矩阵为对角阵.

分析： 先求一组基下的矩阵，然后将此矩阵对角化，从而得到过渡矩阵，由过渡矩阵与此组基得到所求基.

解:

$$E_{11}=\begin{pmatrix}1&0\\0&0\end{pmatrix},E_{12}=\begin{pmatrix}0&1\\0&0\end{pmatrix},E_{21}=\begin{pmatrix}0&0\\1&0\end{pmatrix},E_{22}=\begin{pmatrix}0&0\\0&1\end{pmatrix}.$$

$$T(E_{11},E_{12},E_{21},E_{22})=(E_{11},E_{12},E_{21},E_{22})\begin{pmatrix}0&4&0&0\\1&0&0&0\\0&0&0&4\\0&0&1&0\end{pmatrix}.$$

即
$$A=\begin{pmatrix}0&4&0&0\\1&0&0&0\\0&0&0&4\\0&0&1&0\end{pmatrix}.$$

又 $|\lambda E-A|=(\lambda-2)^2(\lambda+2)^2$.

当$\lambda=2$时，特征向量为

$\xi_1=(2,1,0,0)^T,\xi_2=(0,0,2,1)^T$.

当$\lambda=-2$时，特征向量为

$\xi_3=(2,-1,0,0)^T$, $\xi_4=(0,0,2,-1)^T$.

故

$$P=\begin{pmatrix} 2 & 0 & 2 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 2 & 0 & 2 \\ 0 & 1 & 0 & -1 \end{pmatrix}.$$

从而 $P^{-1}AP=\begin{pmatrix}2&&&\\&2&&\\&&-2&\\&&&-2\end{pmatrix}.$

令 $(X_{11},X_{12},X_{21},X_{22})$

$=(E_{11},E_{12},E_{21},E_{22})P$

$=(2E_{11}+E_{12},2E_{12}+E_{22},2E_{11}-E_{12},2E_{12}-E_{22}).$

$$\Rightarrow X_{11}=\begin{pmatrix}2 & 1\\ 0 & 0\end{pmatrix}, X_{12}=\begin{pmatrix}0 & 0\\ 2 & 1\end{pmatrix},$$

$$X_{21}=\begin{pmatrix}2 & -1\\ 0 & 0\end{pmatrix}, X_{22}=\begin{pmatrix}0 & 0\\ 2 & -1\end{pmatrix}.$$

则T在$X_{11},X_{12},X_{21},X_{22}$下的矩阵为对角阵

$$\begin{pmatrix}2 & & & \\ & 2 & & \\ & & -2 & \\ & & & -2\end{pmatrix}.$$

例7* 已知矩阵 A 与 B 相似，其中

$$A=\begin{bmatrix}1 & -2 & 2\\ -2 & -2 & 4\\ 2 & 4 & -2\end{bmatrix},\quad B=\begin{bmatrix}2 & 0 & 0\\ 0 & -7 & 0\\ 0 & -5 & 2\end{bmatrix}.$$

求可逆阵 P，使 $P^{-1}AP=B$.

解: 由于A、B都与同一个对角阵 Λ相似，因此使用Λ来“过渡”．即存在可逆矩阵P_1、P_2，使

$$P_1^{-1}AP_1=\Lambda,\quad P_2^{-1}BP_2=\Lambda.$$

所以 $(P_1P_2^{-1})^{-1}A(P_1P_2^{-1})=B$. 即

$$P=P_1P_2^{-1}.$$

下面求可逆矩阵 P_1、P_2 .

首先求特征值.

显然 B 的特征值为$\lambda_1=\lambda_2=2,\quad \lambda_3=-7.$

因为 A、B 相似，所以这也是 A 的特征值.

接着求可逆矩阵 P_1.

对于 $\lambda_1=\lambda_2=2$，解$(A-\lambda_1 I)x=0$,

得基础解系$\xi_1=(-2,1,0)^T$，$\xi_2=(2,0,1)^T$.

对于$\lambda_3=-7$，解$(A-\lambda_3 I)x=0$,

得基础解系 $\xi_3=(1,2,-2)^T$.

$$\text{所以}P_1=\left[\xi_1,\xi_2,\xi_3\right]=\begin{bmatrix}-2 & 2 & 1\\ 1 & 0 & 2\\ 0 & 1 & -2\end{bmatrix}.$$

接着求可逆矩阵 P_2.

对于$\lambda_1=\lambda_2=2$，解 $(B-\lambda_1 I)x=0$，

得基础解系 $\tau_1=(1,0,0)^T$， $\tau_2=(0,0,1)^T$.

对于$\lambda_3=-7$ ，解$(B-\lambda_3 I)x=0$,

得基础解系$\tau_3=(0,-9,5)^T$.

所以

$$P_2=\left[\tau_1,\tau_2,\tau_3\right]=\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & -9\\ 0 & 1 & 5\end{bmatrix}.$$

因此

$$P = P_1 P_2^{-1} = \begin{bmatrix} -2 & 2 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} -2 & 2 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 2 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{5}{9} & 1 \\ 0 & -\frac{1}{9} & 0 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 1 & 2 \\ 1 & \frac{-2}{9} & 0 \\ 0 & \frac{7}{9} & 1 \end{bmatrix}.$$

1.2.6 特征值与特征向量

定义1. 2. 4

对n 阶方阵A, 如果存在数域P 中的数λ 和n 维向量α ，使得

$$A(\alpha)=\lambda\alpha,$$

则称λ 为A 的特征值, α为属于特征值λ 的特征向量.称

$$(\lambda E-A)\alpha=0$$

也称为A 的特征方程.

定理1. 2. 8

矩阵$A\in C^{n\times n}$的不同特征值对应的特征向量线性无关.

定义1. 2. 5

矩阵$A \in C^{n\times n}$的主对角线上元素的和

$$a_{11} + a_{22} + \cdots + a_{nn}$$

称为A **的迹**，记作$\mathrm{tr}(A)$.

定理1. 2. 9

矩阵$A \in C^{n\times n}$的特征值分别为$\lambda_1, \lambda_2, \cdots, \lambda_n$，则

(1) 矩阵A的迹等于特征值的和，即

$$a_{11} + a_{22} + \cdots + a_{nn} = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

(2) 矩阵A的的行列式等于特征值的积，即

$$|A| = \lambda_1\lambda_2\cdots\lambda_n.$$

1.2.7 线性变换的特征值与特征向量

定义1. 2. 6

设T 是n 维线性空间V 的一个线性变换, 如果存在数域P 中的数λ 和V 中的元素α , 使得

$$T(\alpha) = \lambda\alpha,$$

则称λ 为T 的特征值, α为属于特征值λ 的特征向量.

定理1. 2. 10

设T 是n 维线性空间V 的一个线性变换，λ 为T 的一个特征值，则特征值λ 对应的全部特征向量加上零向量构成的集合

$$V_\lambda = \{\alpha | T(\alpha) = \lambda\alpha, \alpha \in V\},$$

则V_λ 是V 的一个子空间，称为**特征子空间**.

矩阵论(MATRIX THEORY)

现在讨论线性变换T 的特征值与其在一组基$\alpha_1, \alpha_2, \cdots, \alpha_n$下的矩阵$A$ 的特征值的关系.

设$T(\alpha) = \lambda\alpha$, $\alpha = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n$, 则

$$
\begin{aligned}
T(\alpha) &= T(x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n) \\
&= T(x_1\alpha_1) + T(x_2\alpha_2) + \cdots + T(x_n\alpha_n) \\
&= (T(\alpha_1), T(\alpha_2), \cdots, T(\alpha_n))(x_1, x_2, \cdots, x_n)^T \\
&= (\alpha_1, \alpha_2, \cdots, \alpha_n)A(x_1, x_2, \cdots, x_n)^T,
\end{aligned}
$$

$$
\lambda(\alpha) = (\alpha_1, \alpha_2, \cdots, \alpha_n)\lambda(x_1, x_2, \cdots, x_n)^T,
$$

所以, 有$A(x_1, x_2, \cdots, x_n)^T = \lambda(x_1, x_2, \cdots, x_n)^T$,

即$Ax = \lambda x$.

上式表明：T 的特征值就是矩阵A 的特征值, T 的特征向量在这组基下的坐标就是A 对应的特征向量.由上述定理知, T 在不同基下的矩阵式相似的, 所以其特征值相同, 换言之, 线性变换的特征值与所选的基无关.

矩阵论(MATRIX THEORY)

教材:

李路、王国强、吴中成. 矩阵论及其应用. 上海：东华大学出版社，2019.

矩阵论(MATRIX THEORY)

参考教材:

[1] 戴华. 矩阵论. 北京：科学出版社，2001.

[2] 邸继征. 矩阵论. 北京：科学出版社，2016

[3] 方保镕. 矩阵论（第3版）. 北京：清华大学出版社，2021.

[4] 方保镕. 矩阵论千题习题详解（第2版）. 北京：清华大学出版社，2021.

[5] 张贤达，周杰. 矩阵论及其工程应用. 北京：清华大学出版社，2015.

[6] 张贤达. 矩阵分析与应用（第2版）. 北京：清华大学出版社，2013.

矩阵论(MATRIX THEORY)

参考教材:

[7] F.Z. Zhang．Matrix Theory：Basic Results and Techniques（第2版）．Springer-Verlag New YorkInc., 2011.

[8] R.A. Horn, C.R. Johnson著，张明尧，张凡译. 矩阵分析（第2版）．北京：机械工业出版社，2013.

[9] R.A. Horn, C.R. Johnson. 矩阵分析（英文版.第2版）．北京：人民邮电出版社, 2015.

[10] Gene H. Golub, Charles F. Van Loan著，程晓亮 译. 矩阵计算（第4版）．北京：人民邮电出版社, 2020

[11] Gene H. Golub, Charles F. Van Loan. 矩阵计算（英文版.第4版）．北京：人民邮电出版社, 2014.

[12] John Derbyshire著，张浩 译.代数的历史:人类对未知量的不舍追踪（修订版）．北京：人民邮电出版社, 2021.