现代数学基础丛书 5

多元统计分析引论

张尧庭 方开泰 著

科学出版社

北京

内 容 简 介

本书系统论述多元统计分析的基本理论和方法，力求理论与实际应用并重．只要具有一元统计的知识就可阅读本书．

本书主要内容是：多元正态分布、方差分析、回归分析、因子分析与线性模型、聚类分析和统计量的分布．附录中列出了常用的多元分布表．

读者对象是高等学校数学系教师、高年级学生，应用多元统计的科技工作者．

现代数学基础丛书：5

多元统计分析引论

张尧庭　方开泰　著

责任编辑　刘嘉善

封面设计　黄华斌　陈　敬

出版

北京东黄城根北街16号

邮政编码：100717

印刷

科学出版社发行　各地新华书店经销

*

1982年6月第　一　版　　开本：B5(720×1000)

2003年4月第五次印刷　　印张：30　1/2

印数：15 271—17 770　　字数：487 000

ISBN 7-03-005988-3

定价：60.00元

（如有印装质量问题，我社负责调换〈环伟〉）

序　言

多元统计分析是数理统计学中近二十多年来迅速发展的一个分支.由于电子计算机使用日益广泛,多元分析的方法也很快地应用到各个领域.在国外,从自然科学到社会科学的许多方面,都已证实了多元分析方法是一种很有用的数据处理方法;在我国,多元分析对于地质、气象、水文、国家标准和误差分析等许多方面的研究工作都取得了很大的成绩,引起了广泛的注意.

但是,目前非常缺少系统介绍多元分析的书,现有的书或者过于偏重理论,或者过于偏重单纯地介绍方法.而迫切需要的则是这样一本书,它使得搞数学的人可以从中看到多元分析方法的实际应用,使得搞实际工作的人可以从中看到相应的一些理论.我们正是朝着这个目标来努力的,并希望本书能作为高等学校高年级学生和研究生的入门书,也可以作为实际工作者的参考书.我们假定读者已具备一元统计分析的知识.

全书共分九章,第一章系统介绍多元分析中常用的矩阵知识.本章内容大多只阐述结论,而不给出证明.第二章到第五章,介绍多元正态分布以及常用的方差分析、回归分析和判别分析等方法.第六章、第七章采用比较一般的形式来介绍因子分析和线性模型的内容,读者在熟悉第二章到第五章内容的基础上能更好地理解第六、七两章比较概括抽象的结果.第八章介绍聚类分析的各种典型的方法.第九章专门讨论统计量的分布.最后,为了实际工作的需要,我们在附录里选了六个重要的多元统计表.

本书收集了我国数学工作者的成果,特别是许宝騄先生在多元分析方面的奠基性的成果.

最后,我们感谢参加多元分析讨论班的同志们,他们在讨论中给了我们许多帮助,特别要感谢陈希孺同志,他对本书初稿提出了许多宝贵的意见.

由于水平有限,书中肯定有很多缺点和错误,请读者批评指正.

武汉大学　张尧庭

应用数学研究所　方开泰

1979 年 3 月

目　　录

第一章　矩阵

§1. 线性空间 …… 1

§2. 内积和投影 …… 3

§3. 矩阵的基本性质 …… 5

§4. 分块矩阵的代数运算 …… 12

§5. 特征根及特征向量 …… 16

§6. 对称阵 …… 20

§7. 非负定阵 …… 24

§8. 广义逆 …… 28

§9. 计算方法 …… 36

§10. 矩阵微商 …… 43

§11. 矩阵的标准型 …… 46

§12. 矩阵内积空间 …… 49

第二章　多元正态分布 …… 53

§1. 定义 …… 53

§2. 正态分布的矩 …… 57

§3. 条件分布和独立性 …… 59

§4. 多元正态分布的参数估计 …… 64

§5. μ 和 V 的极大似然估计的性质 …… 69

§6. 多维正态分布的特征 …… 79

§7. 多维正态分布函数的计算 …… 82

§8. 例 …… 91

第三章　样本分布的性质和均值与协差阵的检验 …… 96

§1. 二次型分布 …… 96

§2. 维希特(Wishart)分布 …… 108

§3. 与样本协差阵有关的统计量；T^2 和 Λ 统计量 …… 112

§4. 均值的检验 …… 119

§5. T^2 统计量的优良性 …… 126

§6. 多母体均值的检验 …… 131

§7. 协方差不等时均值的检验 …… 138

§8. 协差阵的检验 …… 140
§9. 独立性检验 …… 149
第四章　判别分析 …… 155
§1. 距离判别 …… 155
§2. 贝叶斯(Bayes)判别 …… 166
§3. 费歇的判别准则 …… 172
§4. 误判概率 …… 179
§5. 附加信息检验 …… 184
§6. 逐步判别 …… 189
§7. 序贯判别 …… 199
第五章　回归分析 …… 203
§1. 问题及模型 …… 203
§2. 最小二乘估计 …… 207
§3. 假设检验 …… 213
§4. 逐步回归 …… 220
§5. 双重筛选逐步回归 …… 230
§6. 回归分析与判别分析的关系 …… 238
第六章　相关 …… 242
§1. 投影 …… 242
§2. 典型相关变量 …… 245
§3. 广义相关系数 …… 252
§4. 主成分分析及主分量分析 …… 257
§5. 因子分析 …… 262
第七章　线性模型 …… 271
§1. 模型 …… 271
§2. 估值 …… 273
§3. 广义线性模型 …… 378
§4. 递推公式 …… 386
§5. 正态线性模型的假设检验 …… 392
§6. 试验设计 …… 302
第八章　聚类分析 …… 314
§1. 相似系数和距离 …… 314
§2. 系统聚类法 …… 320
§3. 系统聚类法的性质 …… 330
§4. 动态聚类法 …… 339

§5. 分解法 …… 349
§6. 有序样品的聚类与预报 …… 352
第九章　统计量的分布 …… 362
§1. 预备知识 …… 362
§2. $I_m(f \mid r_1, \cdots, r_m)$ …… 365
§3. 一元非中心分布 …… 369
§4. Wishart 分布 …… 373
§5. 广义方差的分布 …… 378
§6. 非中心 T^2 分布 …… 382
§7. 样本相关系数的分布 …… 385
§8. $S_1 S_2^{-1}$特征根的联合分布 …… 390
§9. 结束语 …… 402
参考文献 …… 406
附录 …… 409
附表的使用说明 …… 409
表 1　$\Lambda_\alpha(m, n_1, n_2)$表 …… 413
表 2　θ_{max}表 …… 445
表 3　$T^2(m, n)$表 …… 451
表 4　$L(m, \nu)$表 …… 463
表 5　$W(m, n)$表 …… 464
表 6　$M(m, \nu_0, k)$表 …… 473
《现代数学基础丛书》已出版书目 …… 474

第一章　矩　　阵

在这章中我们把矩阵和向量空间中有关多元分析的一些结果介绍给读者,凡是一般教材中有的材料,我们只是罗列一下大都不给出证明.熟悉这些内容的人可以从第二章看起,想看第一章的读者最好把正文中未给证明的结果作为练习,这将为以后各章的阅读带来很多方便.

§1. 线性空间

今后我们限定在实数域 R 上讨论.用 R_n 表示实数域 R 上的全部 n 维向量, R_n 中的一个向量(有时称为元素)a 有 n 个坐标,写成

$$a=\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \text{或}\ a=(a_1,\cdots,a_n)'.$$

1.1 向量的运算　对 R_n 中的向量定义了两种运算:

(i) 加法　设 $a=(a_1,\cdots,a_n)'$, $b=(b_1,\cdots,b_n)'$,则

$$a+b\triangleq^{①}(a_1+b_1,\cdots,a_n+b_n)'.$$

(ii) 数乘　设 c 是一个数, $a=(a_1,\cdots,a_n)'$,则

$$ca\triangleq(ca_1,\cdots,ca_n)'.$$

为了方便,我们用 0 表示坐标全为 0 的向量,这不会引起混淆,从上下文可以判断它是向量还是数.容易证明,对上述定义的加法和数乘,下列关系成立:

(A) 当 a,b,c 均$\in R_n$ 时,则

$$a+b=b+a,(a+b)+c=a+(b+c),$$

$$a+0=a,\quad a+(-a)=0.$$

(B) 当 a,b 均$\in R_n$, c,c_1,c_2 均为实数时,则

$$c(a+b)=ca+cb,(c_1+c_2)a=c_1a+c_2a,$$

$$c_1(c_2a)=c_1c_2a,\quad 1\cdot a=a.$$

我们称 R_n 是实数域 R 上的线性空间或向量空间.如果 R_n 中的子集 L,对加法和

---

① 表示该符号左边的内容由右边的符号表示,也可以理解为定义.

数乘这两种运算是封闭的(即运算的结果仍在 L 中),且 L 不是空集,则称 L 是 R_n 中的一个子空间.例如

$$L=\{(a,\underbrace{0,\cdots,0}_{n-1\text{个}})':a\in R\}$$

就是 R_n 的一个子空间.

1.2 线性相关　线性空间最重要的概念就是线性相关与线性无关.

如有不全为 0 的一组数 $c_1,\cdots,c_k$ 使 $c_1a_1+\cdots+c_ka_k=0$,则称向量 $a_1,\cdots,a_k$是线性相关的;否则,就称 $a_1,\cdots,a_k$ 是线性无关的.从这个定义可以看出:

(i) 任何含有 0 向量的向量集总是线性相关的;

(ii) $a_1,\cdots,a_k$ 线性无关的充要条件是:如果一组数 $c_1,\cdots,c_k$ 使 $c_1a_1+\cdots+c_ka_k=0$,则 $c_1=\cdots=c_k=0$;

(iii) 设 $a_1,\cdots,a_k$ 是非零向量,它们线性相关的充要条件是:存在 i 使 $a_i=b_1a_1+b_2a_2+\cdots+b_{i-1}a_{i-1}+b_{i+1}a_{i+1}+\cdots+b_ka_k$,其中 $b_1,\cdots,b_k$ 是一组实数.也即存在 i 使 a_i 是其他向量 $a_1,\cdots,a_{i-1},a_{i+1},\cdots,a_k$ 的线性组合.

给定了 R_n 中某些向量 $a_1,\cdots,a_k$,考虑由这些向量所有可能的线性组合 $\sum\limits_{i=1}^{k}c_ia_i$ 组成的集合:

$$\mathscr{L}(a_1,\cdots,a_k)=\left\{\sum_{i=1}^{k}c_ia_i:c_1,\cdots,c_k\text{ 均为实数}\right\},$$

它显然对加法和数乘这两种运算是封闭的,它是一个子空间,称它是由向量 $a_1,\cdots,a_k$ 生成的子空间.

1.3 基　设 $\mathscr{L}$是 R_n 中的一个子空间,如果存在 $a_1,\cdots,a_k$ 使 $\mathscr{L}=\mathscr{L}(a_1,\cdots,a_k)$,且 $a_1,\cdots,a_k$ 线性无关,则称 $a_1,\cdots,a_k$ 是 $\mathscr{L}$的一组基.

可以证明子空间 $\mathscr{L}$中如有两组基,那么这两组基中向量的个数一定相同,因此我们把子空间 $\mathscr{L}$中一组基所含的向量的个数称为 $\mathscr{L}$的维数.例如

$$e_i=(\underbrace{0,0,\cdots,0}_{i-1\text{个}},1,\underbrace{0,\cdots,0}_{n-i\text{个}})',\ i=1,2,\cdots,n$$

显然是线性无关的,并且 $R_n=\mathscr{L}(e_1,\cdots,e_n)$,因此线性空间 R_n 的维数是 n.有时就称 R_n 为 n 维线性空间.很显然,$\mathscr{L}(e_1),\cdots,\mathscr{L}(e_n)$都是一维的子空间;$\mathscr{L}(e_1,e_2),\cdots,\mathscr{L}(e_{n-1},e_n)$等都是二维的子空间;….

从基的定义立即可知,如果 $a_1,\cdots,a_k$ 是子空间 $\mathscr{L}$的一组基,那么 $\mathscr{L}$中任一向量 a 都可被 $a_1,\cdots,a_k$ 的线性组合来表示,而且这种表示法是唯一的.

1.4 直接和　设 $\mathscr{L}$是一个子空间,如果有 k 个子空间 $\mathscr{L}_1,\cdots,\mathscr{L}_k$ 使得对每一 $a\in\mathscr{L}$,能唯一地表示为 $a_1+a_2+\cdots+a_k$,其中 $a_i\in\mathscr{L}_i$, $i=1,\cdots,k$,则称 $\mathscr{L}$是$\mathscr{L}_1,\cdots,\mathscr{L}_k$ 的直接和,记为 $\mathscr{L}=\mathscr{L}_1+\cdots+\mathscr{L}_k$.沿用 1.3 中 e_i 的定义,可以看出

$$R_n = \mathscr{L}(e_1) + \cdots + \mathscr{L}(e_n).$$

§2. 内积和投影

R_n 中任给两个向量 $a=(a_1,\cdots,a_n)'$, $b=(b_1,\cdots,b_n)'$,定义 a,b 的内积 $(a,b) \triangleq \sum_{i=1}^{n} a_i b_i$.易见内积(a,b)满足下列性质:

(i) $(a,b)=(b,a)$;

(ii) $(a,a)\geqslant 0$;$(a,a)=0 \Leftrightarrow a=0$;

(iii) $(ca,b)=(a,cb)=c(a,b)$对一切 $c\in R$ 成立;

(iv) $(a,h+g)=(a,h)+(a,g)$,

$$(h+g,b)=(h,b)+(g,b).$$

我们把(a,a)的算术平方根称为 a 的长度,记作 $\|a\|$.容易验证

$$|(a,b)|^2 \leqslant \|a\|^2 \|b\|^2 \quad \text{(Schwarz 不等式)}, \tag{2.1}$$

$$\|a+b\| \leqslant \|a\| + \|b\| \quad \text{(三角不等式)}. \tag{2.2}$$

2.1 标准正交基 如果 R_n 中的子空间 $\mathscr{L}$的基 $a_1,\cdots,a_k$ 具有性质:

$$(a_i,a_i)=1, i=1,2,\cdots,k,$$

$$(a_i,a_j)=0, i\neq j, i,j=1,2,\cdots,k,$$

则称 $a_1,\cdots,a_k$ 是$\mathscr{L}$的一组标准正交基,因为我们把 $\|a\|=1$ 的向量称为标准化的向量,有时也称为单位向量(指它的长度是 1,以它作单位).当$(a,b)=0$ 时,我们就称 a 与 b 正交,$(a_i,a_j)=0$ 表示 a_i 与 a_j 正交.引入克劳涅克尔的 δ 符号:

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

则 $a_1,\cdots,a_k$ 是$\mathscr{L}$的一组标准正交基$\Leftrightarrow a_1,\cdots,a_k$ 均属于$\mathscr{L}$,且$(a_i,a_j)=\delta_{ij}$, $i,j=1,2,\cdots,k$.

2.2 投影 在 R_n 中,给定一个向量 a 及子空间$\mathscr{L}$,就可考虑 a 在$\mathscr{L}$中的投影.如果在$\mathscr{L}$中存在 b 使 $\|a-b\|=\inf\limits_{x\in\mathscr{L}}\|a-x\|$,则称 b 是 a 在$\mathscr{L}$中的投影.可以证明投影是存在而且唯一的.下面先证明一条关于投影的重要性质:

b 是 a 在$\mathscr{L}$中的投影$\Leftrightarrow(a-b,x)=0$ 对一切 $x\in\mathscr{L}$成立.

证明"$\Rightarrow$",若有 $x\in\mathscr{L}$使$(a-b,x)\neq 0$,由于对一切 λ 有

$$\begin{aligned}(a-b,a-b) &\leqslant (a-b+\lambda x, a-b+\lambda x)\\ &=(a-b,a-b)+2\lambda(a-b,x)+\lambda^2(x,x),\end{aligned}$$

只要选 $\lambda=-\varepsilon(a-b,x)$, $\varepsilon>0$,上式就可写成

$$(a-b,x)^2(\varepsilon^2\|x\|^2-2\varepsilon)\geqslant 0$$

对一切 $\varepsilon>0$ 成立,而这是不可能的,因而$(a-b,x)=0$.

"$\Leftarrow$",由于$(a-x,a-x)=(a-b+b-x,a-b+b-x)$

$$=(a-b,a-b)+2(a-b,b-x)+(b-x,b-x),$$

注意 $b\in\mathscr{L}$, $x\in\mathscr{L}$,就有 $b-x\in\mathscr{L}$,因而$(a-b,b-x)=0$,于是由

$$(a-x,a-x)=(a-b,a-b)+(b-x,b-x)$$

看出 $x=b$ 时使 $\|a-x\|$ 达到最小值.

从这个证明就可看出投影存在时,只有一个,投影的存在性在 2.4 中给出.

2.3 格兰姆-施密特正交化方法　一组向量 $a_1,\cdots,a_k$ 是线性无关的,则可依照下述方法求出子空间 $\mathscr{L}(a_1,\cdots,a_k)$的一组标准正交基.

先取 $b_1=a_1$,因为 $a_1,\cdots,a_k$ 线性无关,因此 $a_1\neq 0$,即$(a_1,a_1)\neq 0$;然后选 $b_2=a_2-h_{21}b_1$ 使$(b_2,b_1)=0$,即选 h_{21} 使$(a_2-h_{21}b_1,b_1)=0$,即 $h_{21}=\dfrac{(a_2,b_1)}{(b_1,b_1)}=\dfrac{(a_2,a_1)}{(a_1,a_1)}$;再选 $b_3=a_3-h_{32}b_2-h_{31}b_1$ 使$(b_3,b_2)=(b_3,b_1)=0$,定出系数 h_{32}, h_{31};$\cdots$;一般地,选

$$b_i=a_i-h_{ii-1}b_{i-1}-\cdots-h_{i1}b_1$$

使$(b_i,b_{i-1})=(b_i,b_{i-2})=\cdots=(b_i,b_1)=0$,定出系数 $h_{ii-1},\cdots,h_{i1}$;$\cdots$,这样就求出一组向量 $b_1,\cdots,b_k$.容易看出,如此求得的 $b_1,\cdots,b_k$ 是两两正交的,令

$$\beta_i=\frac{1}{\|b_i\|}b_i,\ i=1,2,\cdots,k,$$

则 $\beta_1,\cdots,\beta_k$ 就是标准正交的向量组.只要证明一组不含零向量的两两正交的向量一定是线性无关的,那么由 $\beta_1,\cdots,\beta_k$ 是标准正交的就可知它们确实是一组基.若不然,$\beta_1,\cdots,\beta_k$ 线性相关,存在 $c_1,\cdots,c_k$ 不全为 0,使 $\sum\limits_{i=1}^{k}c_i\beta_i=0$, 因此

$$0=\left(\sum_i c_i\beta_i,\sum_i c_i\beta_i\right)=\sum_i\sum_j c_ic_j(\beta_i,\beta_j)=\sum_{i=1}^{k}c_i^2,$$

这就导出 $c_1=\cdots=c_k=0$,矛盾.

格兰姆-施密特正交化方法是一种典型的方法,今后我们将不止一次地引用它.

2.4 正交和　设 $\mathscr{L}$是一个子空间,$\mathscr{L}_1,\cdots,\mathscr{L}_k$ 也是一些子空间,$\mathscr{L}$是 $\mathscr{L}_1,\cdots,\mathscr{L}_k$ 的直接和,并且当 $i\neq j$ 时,只要 $a_i\in\mathscr{L}_i$, $a_j\in\mathscr{L}_j$,则$(a_i,a_j)=0$,那么我们就说 $\mathscr{L}$是 $\mathscr{L}_1,\cdots,\mathscr{L}_k$ 的正交和,记为

$$\mathscr{L}=\mathscr{L}_1\dot{+}\mathscr{L}_2\dot{+}\cdots\dot{+}\mathscr{L}_k.$$

一般说来,如果 $\beta_1,\cdots,\beta_n$ 是 R_n 中的一组基,那就一定有 $R_n=\mathscr{L}(\beta_1)+\cdots+\mathscr{L}(\beta_n)$,但是 $\beta_1,\cdots,\beta_n$ 不一定是彼此正交的,因此 R_n 不一定是 $\mathscr{L}(\beta_1),\cdots,\mathscr{L}(\beta_n)$的正交

和，只有当 $\beta_1,\cdots,\beta_n$ 是正交基时，$R_n=\mathscr{L}(\beta_1)\dot{+}\cdots\dot{+}\mathscr{L}(\beta_n)$. 取 $e_i=(\underbrace{0,0,\cdots,0}_{i-1\text{个}},1,\underbrace{0,\cdots,0}_{n-i\text{个}})'$, $i=1,2,\cdots,n$ 后，显然有

$$R_n=\mathscr{L}(e_1)\dot{+}\mathscr{L}(e_2)\dot{+}\cdots\dot{+}\mathscr{L}(e_n).$$

特别地，如果 $R_n=\mathscr{L}_1\dot{+}\mathscr{L}_2$，则称 $\mathscr{L}_1$ 是 $\mathscr{L}_2$ 的正交补空间，或者称 $\mathscr{L}_2$ 是 $\mathscr{L}_1$ 的正交补空间. 子空间 $\mathscr{L}$ 的正交补空间用 $\mathscr{L}^{\perp}$ 表示，因此总有

$$\mathscr{L}\dot{+}\mathscr{L}^{\perp}=R_n.$$

正交补空间和投影有密切的关系. 因为 $R=\mathscr{L}\dot{+}\mathscr{L}^{\perp}$，因此每一个 R 中的向量 x 总可以写成 u_x+v_x，其 $u_x\in\mathscr{L}$, $v_x\in\mathscr{L}^{\perp}$. 由于 $x-a=x-u_x+u_x-a$，因而就有

$$\begin{aligned}(x-a,x-a)=&(x-u_x,x-u_x)+(u_x-a,u_x-a)\\&+2(x-u_x,u_x-a).\end{aligned}$$

注意到 $x-u_x=v_x\in\mathscr{L}^{\perp}$，如果 $a\in\mathscr{L}$ 时，就有 $u_x-a\in\mathscr{L}$，因此 $(x-u_x,u_x-a)=0$. 这就告诉我们

$$(x-a,x-a)\geqslant(x-u_x,x-u_x)$$

对一切 $a\in\mathscr{L}$ 成立，且等号成立的充要条件是 $u_x-a=0$，也即 $a=u_x$. 这就告诉我们对给定的子空间 $\mathscr{L}$, R_n 中任一向量 x 在 $\mathscr{L}$ 中的投影是存在的.

§3. 矩阵的基本性质

将 mn 个实数按一定顺序排成的阵

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

称为大小是 $m\times n$ 的矩阵，它有 m 行、n 列. 当 $m=n$ 时就称为方阵. 今后用大写的拉丁字母表示矩阵，相应的小写字母附有双重足标的就表示矩阵中的元素，例如 a_{ij} 就是矩阵 A 中第 i 行第 j 列的元素，因此有时也写成 $\underset{m\times n}{A}=(a_{ij})_{m\times n}$，或简写为 $A=(a_{ij})$.

3.1 代数运算　对矩阵定义如下的运算：

加法　设 A,B 都是 $m\times n$ 的阵，则

$$A+B\triangleq(a_{ij}+b_{ij})，称为\ A,B\ 的和.$$

数乘　$cA\triangleq(ca_{ij})$.

乘法　设 A 是 $m\times n$ 的阵，B 是 $n\times p$ 的阵，则

$$AB \triangleq \left(\sum_{\alpha=1}^{n} a_{i\alpha}b_{\alpha j}\right) \text{是 } m\times p \text{ 的阵,称为 } A, B \text{ 的乘积}.$$

容易证明这些运算具有下列性质：

$$A+B=B+A, \qquad (A+B)+C=A+(B+C),$$
$$(c+d)A=cA+dA, \quad c(A+B)=cA+cB,$$
$$(AB)C=A(BC), \qquad A(B+C)=AB+AC,$$
$$(A+B)C=AC+BC.$$

注意,矩阵相加时,大小一定要相同;矩阵 A, B 相乘时, A 的列数必须和 B 的行数相同,即 A 是 $m\times n$ 阵时, B 一定是 $n\times p$ 的阵才可相乘. AB 有意义时, BA 不一定有意义,即使 AB 与 BA 都有意义,但大小还不相同,即使大小相同,也不一定有等式 $AB=BA$,这些都是需要注意的.

向量可以看成只有一列的矩阵,矩阵 $\underset{n\times m}{A}$ 也可看成是 m 个 n 维的向量按一定顺序排成的阵,记 $\underset{n\times m}{A}=(a_1\cdots a_m)$,则 a_i 就是 A 中第 i 列组成的向量.这一种向量和矩阵的联系今后是经常要用的.今后把 A'中的列向量称为 A 的行向量.

元素全为 0 的矩阵记作 O,元素全为 1 的矩阵记作 J,元素全为 1 的向量记作 $\mathbf{1}$,它们的大小视上下文而定.矩阵

$$\begin{pmatrix} 1 & & & O \\ & 1 & & \\ & & \ddots & \\ O & & & 1 \end{pmatrix}_{n\times n}$$

是 $n\times n$ 的阵,它的主对角元素全为 1,其余全为 0,称它为单位阵,记作 I.实际上 $I=(\delta_{ij})$. I 的大小不标明时由上下文来定.易见

$$A+0=A, \ \underset{n\times m}{A} I_m = I_n \underset{n\times m}{A} = A.$$

3.2 转置和逆　在矩阵运算中还有两种运算是常见的：

转置:把 A 的行改成列,列改成行后得到的矩阵,记作 A'.例如 $a=\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$,则 $a'=(a_1\cdots a_n)$.如果

$$\underset{m\times n}{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

则

$$\underset{n\times m}{A'}=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

注意到向量 $a=\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ 是 $n\times 1$ 的阵,则向量 $b=\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ 也是 $n\times 1$ 的阵,于是按矩阵乘法的定义就得:

$$\underset{1\times n}{a'}\underset{n\times 1}{b}=(a_1\cdots a_n)\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}=\sum_{i=1}^{n} a_i b_i,(\text{就是 } a,b \text{ 的内积}) \tag{3.1}$$

$$\underset{n\times 1}{a}\underset{1\times n}{b'}=\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}(b_1\cdots b_n)=\begin{pmatrix} a_1 b_1 & a_1 b_2 & \cdots & a_1 b_n \\ a_2 b_1 & a_2 b_2 & \cdots & a_2 b_n \\ \vdots & \vdots & & \vdots \\ a_n b_1 & a_n b_2 & \cdots & a_n b_n \end{pmatrix} \tag{3.2}$$

它是一个 $n\times n$ 的矩阵.

如果 A 是 $m\times n$ 的阵,将 A 写成 $\begin{pmatrix} a_1' \\ \vdots \\ a_m' \end{pmatrix}$,则 $a_1,\cdots,a_m$ 是 m 个 n 维的向量,$a_i'=(a_{i1}\ a_{i2}\cdots a_{in})$;如果 B 是 $n\times p$ 的阵,将 B 写成 $(b_1\cdots b_p)$,则 b_i 都是 n 维向量,$b_i=\begin{pmatrix} b_{1i} \\ \vdots \\ b_{ni} \end{pmatrix}$.于是

$$AB=\begin{pmatrix} a_1' \\ \vdots \\ a_m' \end{pmatrix}(b_1\cdots b_p)=\begin{pmatrix} a_1' b_1 & \cdots & a_1' b_p \\ \vdots & \ddots & \vdots \\ a_m' b_1 & \cdots & a_m' b_p \end{pmatrix},$$

也即乘积矩阵 AB 中(i,j)位置的元素 $\sum_{\alpha=1}^{n} a_{i\alpha} b_{\alpha j}$ 是 A 中第 i 行向量 a_i 与 B 中第 j 列向量 b_j 的内积.

逆:如果 A 是方阵,且有 B 使

$$AB=BA=I,$$

则称 B 是 A 的逆,记作 A^{-1}.给定了方阵 A,A^{-1}可能不存在,但它存在时一定唯一.我们称存在逆阵的 A 为非奇异阵,不存在逆阵的 A 是奇异阵.易见有公式:

$$\begin{cases}(A')' = A,(AB)' = B'A', \\ (A^{-1})^{-1} = A,(A')^{-1} = (A^{-1})',(AB)^{-1} = B^{-1}A^{-1}.\end{cases} \tag{3.3}$$

3.3 秩 为了方便,今后规定,不加声明时矩阵 A 的元素用 a_{ij}表示,A 的列向量用 a_i 表示,A 的行向量用 $a_{(i)}$表示,即

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm}\end{bmatrix} = (a_1 \cdots a_m) = \begin{bmatrix} a'_{(1)} \\ \vdots \\ a'_{(n)}\end{bmatrix},$$

$$a_i = (a_{1i} a_{2i} \cdots a_{ni})', a_{(i)} = (a_{i1} a_{i2} \cdots a_{im})'.$$

对给定的矩阵 $\underset{n\times m}{A}$,$\mathscr{L}(a_1,\cdots,a_m)$的维数称为矩阵 A 的秩,记作 rk A 或 rk(A).即 A 的秩就是 A 的列所生成的子空间的维数.可以证明对给定的 $\underset{n\times m}{A}$,$\mathscr{L}(a_1,\cdots,a_m)$与 $\mathscr{L}(a_{(1)},\cdots,a_{(n)})$的维数相同,因此就有 rk$(A)$=rk$(A')$.今后用 $\mathscr{L}(A)$表示 $\mathscr{L}(a_1,\cdots,a_m)$,自然 $\mathscr{L}(a_{(1)},\cdots,a_{(n)})$也就是 $\mathscr{L}(A')$.实际上

$$\mathscr{L}(\underset{n\times m}{A}) = \{\underset{n\times m}{A}\ \underset{m\times 1}{x}: \quad x \in R_m\}.$$

由于 $\mathscr{L}(\underset{n\times m}{A}\underset{m\times l}{B}) = \{ABu: u\in R_l\}\subset\mathscr{L}(A)$,因此就有

$$\mathrm{rk}(AB) \leqslant \mathrm{rk}(A).$$

又由于 rk(AB)=rk$((AB)')$=rk$(B'A')\leqslant$rk(B')=rk(B),这样就证明了

$$\mathrm{rk}(AB) \leqslant \min(\mathrm{rk}(A),\mathrm{rk}(B)).$$

这就告诉我们矩阵乘积的秩不超过因子的秩.

如果 $B=PAQ$,且 P^{-1},Q^{-1}存在,于是 $A=P^{-1}BQ^{-1}$,这就可导出 rk(A)=rk(B),也即 A 左乘或右乘非奇异阵时,不会改变 A 的秩.

一般说来 $\mathscr{L}(A)$与 $\mathscr{L}(AB)$是不等的,$\mathscr{L}(AB)\subset\mathscr{L}(A)$.但当 rk$(AB)$=rk$(A)$时,这两个子空间的维数相同,一个又在另一个之内,它们就一定相等,也即 $\mathscr{L}(AB)=\mathscr{L}(A)$,也即存在 C 使 $A=ABC$.于是 $x'A=0\Rightarrow x'AB=0\Rightarrow x'ABC=0\Rightarrow x'A=0$.即有

$$x'A = 0 \Leftrightarrow x'AB = 0.$$

这就告诉我们:如果 rk(A)=rk(AB),则由 $C_1AB=C_2AB$ 可导出 $C_1A=C_2A$(即两边可以“消去”矩阵 B).这是因为 $C_1AB=C_2AB\Leftrightarrow(C_1-C_2)AB=0\Leftrightarrow(C_1-C_2)A=0\Leftrightarrow C_1A=C_2A$.同理,如果 rk$(AB)$=rk$(B)$,则由 $ABC_1=ABC_2$ 可导出 $BC_1=BC_2$,即可“消去”A.

3.4 行列式 给定一个 $n\times n$ 的方阵 A,定义 A 的行列式$|A|$如下:

$$|A| = \sum_{\sigma}(-1)^{\sigma(i_1,i_2,\cdots,i_n)} a_{1i_1}a_{2i_2}\cdots a_{ni_n},$$

其中

$$\sigma(i_1,\cdots,i_n)=\begin{cases}1, 当(i_1,\cdots,i_n)是奇排列,\\ 0, 当(i_1,\cdots,i_n)是偶排列,\end{cases}$$

$\sum\limits_{\sigma}$ 是对一切排列$(i_1,\cdots,i_n)$求和.

关于行列式的性质和有关的证明,这里不逐一罗列了,这些内容在一般代数书中都有,下面只列出几条常用的.

将方阵 A 中删去第 i 行、第 j 列的元素后所留下的子阵相应的行列式,称为 A 中元素 a_{ij}的余子式.将 a_{ij}的余子式乘以$(-1)^{i+j}$后就称为 a_{ij}的代数余子式,记作 A_{ij}.可以证明:

$$\begin{cases}\text{(i)}\ |A'|=|A|=\sum\limits_{i=1}^{n}a_{ij}A_{ij}=\sum\limits_{j=1}^{n}a_{ij}A_{ij},\ i,j=1,\cdots,n;\\ \text{(ii)}\sum\limits_{i=1}^{n}a_{ij}A_{il}=0, 当\ j\neq l,\ j,l=1,\cdots,n;\\ \text{(iii)}\ 当\ |A|\neq 0\ 时,\\ \qquad |A|\,A^{-1}=(A_{ij})';\\ \text{(iv)}\ |AB|=|A|\,|B|;\\ \text{(v)}\ |\underset{n\times n}{A}|\neq 0\Leftrightarrow \underset{n\times n}{A}\ 非奇异\Leftrightarrow \mathrm{rk}\,A=n;\\ \text{(vi)}\ |A^{-1}|=|A|^{-1}.\end{cases}\tag{3.4}$$

3.5 线性方程组 设有 m 个未知数 $x_1,\cdots,x_m$ 及 n 个方程:

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1m}x_m=b_1,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2m}x_m=b_2,\\ \cdots\cdots\cdots\cdots\\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nm}x_m=b_n,\end{cases}$$

用矩阵的形式来写,记

$$A=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1m}\\ a_{21}&a_{22}&\cdots&a_{2m}\\ \vdots&\vdots&&\vdots\\ a_{n1}&a_{n2}&\cdots&a_{nm}\end{bmatrix},\quad x=\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_m\end{bmatrix},\quad b=\begin{bmatrix}b_1\\ b_2\\ \vdots\\ b_n\end{bmatrix}.$$

那么,上述线性方程组可写为

$$Ax=b.$$

考虑 A 的列向量生成的线性空间 $\mathscr{L}(A)$,易见 $Ax=b$ 有解的充要条件是 $b\in\mathscr{L}(A)$,即 $\mathscr{L}(A)=\mathscr{L}(A,b)$.通常把矩阵$(Ab)$称为线性方程组 $Ax=b$ 的增广矩阵,A 称为系数矩阵.于是就有:

$$Ax=b\ 有解\Leftrightarrow 系数矩阵的秩等于增广矩阵的秩$$

$$\Leftrightarrow \mathrm{rk}(A) = \mathrm{rk}(Ab).$$

如果 $b=0$,线性方程组 $Ax=0$ 称为齐次方程,很明显,如果 A 的列向量 $a_1,\cdots,a_m$ 是线性无关的,则 $Ax=0$ 只有 $x=0$ 这一个解;反之也真,$Ax=0$ 只有 $x=0$ 这个解,$a_1,\cdots,a_m$ 就是线性无关的向量.因此,就得

$$\underset{n\times m}{A}x = 0 \text{ 有非零解(即存在 } x\neq 0 \text{ 使 } Ax=0)$$
$$\Leftrightarrow a_1,\cdots,a_m \text{ 线性相关} \Leftrightarrow \mathscr{L}(A) \text{ 的维数小于 } m$$
$$\Leftrightarrow \mathrm{rk}(\underset{n\times m}{A}) < m.$$

如果 A 是一个 $n\times n$ 方阵,则 $Ax=b$ 有解的充分条件是 A^{-1} 存在,并且 $x=A^{-1}b$ 就是解.注意到 A^{-1} 存在时,

$$A^{-1} = \frac{1}{|A|}(A_{ij})' = \frac{1}{|A|}\begin{pmatrix} A_{11} & \cdots & A_{n1} \\ A_{12} & \cdots & A_{n2} \\ \vdots & & \vdots \\ A_{1n} & \cdots & A_{nn} \end{pmatrix}, \tag{3.5}$$

其中 A_{ij} 是元素 a_{ij} 的代数余子式,那么 $Ax=b$ 的解 x 可以用 A 的行列式及代数余子式 A_{ij} 和 b 表出,很明显,此时解存在而且唯一.但是 A^{-1} 不存在时,$Ax=b$ 仍然可能有解,只要 b 属于 $\mathscr{L}(A)$,也即 b 能由 A 的列向量 $a_1,\cdots,a_n$(此时有 n 列)线性表示.设 $b=\sum_{i=1}^{n} c_i a_i$,则 $(c_1,c_2,\cdots,c_n)'$ 就是方程 $Ax=b$ 的解.如果 $b=0$,A^{-1} 存在时,$Ax=0$ 只有 $x=0$ 的解;反之,如果 $Ax=0$ 只有 $x=0$ 的解,则 $a_1,\cdots,a_n$ 线性无关,因此 $\mathrm{rk}(A)=n$,于是 A^{-1} 存在.这就告诉我们:当 A 是方阵时,$Ax=0$ 有非零解的充要条件是 $|A|=0$.

今后为了方便,线性方程组 $Ax=b$ 有解时,就说 $Ax=b$ 是相容的,$Ax=b$ 无解时,就说 $Ax=b$ 是不相容的,或矛盾的.

3.6 初等变换　在矩阵运算中,常常会用到初等变换,设 A 是 $n\times m$ 的矩阵,按 A 的列写出,$A=(a_1\cdots a_m)$,考虑下列变换:

(i) 将指定的某二列互换,例如第 1,2 列互换,即将 $A\to(a_2 \quad a_1 \quad a_3\cdots a_m)$;

(ii) 将指定某一列均乘以实数 c,例如第 1 列乘以实数 c,即将 $A\to(ca_1 \quad a_2 \quad a_3\cdots a_m)$;

(iii) 将第 i 列的元素逐个加上第 j 列相应元素的 c 倍,即将 $A\to(a_1 \quad a_2\cdots a_i+ca_j \quad a_{i+1}\cdots a_m)$.这些变换称为对 A 施行“列初等变换”,这些变换都是对 A 的列向量进行的,由矩阵的乘法知道,上述变换相当于对 A 阵右乘某些矩阵.将第 i 列与第 j 列互换,相当于右乘

$$
\underset{m\times m}{P_{ij}} = \begin{array}{cc}
\left(\begin{array}{ccccccccccc}
1 &&&&&&&&&& \\
& \ddots &&&&&&&&& \\
&& 1 &&&&&&&& \\
&&& 0 && \cdots && 1 &&& \\
&&& \vdots & 1 &&& \vdots &&& \\
&&& \vdots && \ddots && \vdots &&& \\
&&& \vdots &&& 1 & \vdots &&& \\
&&& 1 && \cdots && 0 &&& \\
&&&&&&&& 1 && \\
&&&&&&&&& \ddots & \\
&&&&&&&&&& 1
\end{array}\right) & \begin{array}{c} \\ \\ i \\ \\ \\ \\ \\ \\ j \\ \\ \\ \end{array} \\
\begin{array}{ccccccccccc} & & & i & & & & j & & & \end{array} &
\end{array}.
$$

P_{ij}的元素

$$
p_{\alpha\beta} = \begin{cases} 1, \text{当 } \alpha = \beta, \alpha \neq i, j; \text{或 } \alpha = i, \beta = j; \text{或 } \alpha = j, \beta = i; \\ 0, \text{其他情况}, i, j = 1, 2, \cdots, m. \end{cases}
$$

形如 P_{ij}的矩阵称为置换阵,它是对单位阵 I 作第 i 列与第 j 列的置换得来的.将第 i 列乘以 c,相当于右乘

$$
\begin{array}{cc}
\left(\begin{array}{ccccccc}
1 &&&&&& \\
& \ddots &&& O && \\
&& 1 &&&& \\
&&& c &&& \\
&&&& 1 && \\
&& O &&& \ddots & \\
&&&&&& 1
\end{array}\right) & \begin{array}{c} \\ \\ \\ i \\ \\ \\ \\ \end{array}
\end{array},
$$

它是对单位阵第 i 列乘以 c 所得的阵.变换(iii)相当于右乘

$$
\begin{array}{cc}
\left(\begin{array}{ccccccc}
1 &&&&&& \\
& \ddots &&&&& \\
&& 1 & \cdots & c && \\
&&& \ddots &&& \\
&& O && 1 && \\
&&&&& \ddots & \\
&&&&&& 1
\end{array}\right) & \begin{array}{c} \\ \\ j \\ \\ i \\ \\ \\ \end{array} \\
\begin{array}{ccccccc} & & j & & i & & \end{array} &
\end{array}, \quad j < i,
$$

或

$$
\begin{array}{c@{}c}
\begin{pmatrix}
1 & & & & & & \\
& \ddots & & & & & \\
& & 1 & O & & & \\
& & \vdots & \ddots & & & \\
& & c & \cdots & 1 & & \\
& & & & & \ddots & \\
& & & & & & 1
\end{pmatrix} & \begin{matrix} \\ \\ i \\ \\ j \\ \\ \\ \end{matrix} \\
\begin{matrix} \quad\quad\quad i & \quad\quad j \quad\quad\quad\quad\quad \end{matrix} &
\end{array}, \quad i < j,
$$

它们也是对单位阵 I 作相应的变换得来的.显然,这些矩阵都是非奇异的,如果对 A 的行施行相当于(i),(ii),(iii)的变换,则就等于对 A 左乘上述这一类矩阵(注意大小不同).这些矩阵就称为初等矩阵;对 A 左乘或右乘初等矩阵,就称对 A 进行初等变换.下面要考察经过初等变换后,矩阵 A 会变成什么形状.讨论这些问题时,用分块矩阵的运算比较方便,因此放在§4.在§9 介绍一些算法时,还要考虑初等变换,希望读者注意这一点.

§4. 分块矩阵的代数运算

在矩阵运算时,往往需要将矩阵分块,我们引入一些分块的记号.设

$$\underset{n\times m}{A} = (a_{ij}),\ i = 1,\cdots,n,\ j = 1,\cdots,m.$$

记

$$
\begin{aligned}
A_{11} &= (a_{ij}),\ i = 1,\cdots,r,\ j = 1,\cdots,s,\\
A_{12} &= (a_{ij}),\ i = 1,\cdots,r,\ j = s+1,\cdots,m,\\
A_{21} &= (a_{ij}),\ i = r+1,\cdots,n,\ j = 1,\cdots,s,\\
A_{22} &= (a_{ij}),\ i = r+1,\cdots,n,\ j = s+1,\cdots,m.
\end{aligned}
$$

则 A 可写成

$$
A = \begin{array}{c@{}c}
\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} & \begin{matrix} r \\ n-r \end{matrix} \\
\begin{matrix} \quad s & \quad\quad m-s \end{matrix} &
\end{array}.
$$

我们约定,如果 $r, s, n-r, m-s$ 中有一个是 0,则相应的子块就不出现.有时为了使分块和乘积不致发生混淆,我们用

$$
A = \left(\begin{array}{c:c} A_{11} & A_{12} \\ \hdashline A_{21} & A_{22} \end{array}\right)
$$

表示,希望读者注意.

4.1 分块的运算　设 A, B 是大小相同的两个矩阵,并且分块的大小也一致,即

$$A=\begin{array}{cl}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} & \begin{array}{l} r \\ n-r \end{array} \\ \begin{array}{cc} \;s\; & \;m-s\; \end{array} & \end{array},\quad B=\begin{array}{cl}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} & \begin{array}{l} r \\ n-r \end{array} \\ \begin{array}{cc} \;s\; & \;m-s\; \end{array} & \end{array},$$

则有

(i) $A+B=\begin{pmatrix} A_{11}+B_{11} & A_{12}+B_{12} \\ A_{21}+B_{21} & A_{22}+B_{22} \end{pmatrix}$；

(ii) $cA=\begin{pmatrix} cA_{11} & cA_{12} \\ cA_{21} & cA_{22} \end{pmatrix}$对一切实数 c 成立；

(iii) $A'=\begin{pmatrix} A'_{11} & A'_{21} \\ A'_{12} & A'_{22} \end{pmatrix}$.

4.2 分块的逆 分块求逆是矩阵运算中很有用的一种方法，由它可以演变出一些矩阵求逆的公式，这一小节第一次介绍这个内容，以后还会多次重复出现，请读者注意这一方法.

设 A,B 是可乘的两个阵，将 A,B 相应地分块，使得分块相应的乘积有意义.即若

$$\underset{n\times m}{A}=\begin{array}{cl}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} & \begin{array}{l} r \\ n-r \end{array} \\ \begin{array}{cc} \;s\; & \;m-s\; \end{array} & \end{array},\quad \underset{m\times l}{B}=\begin{array}{cl}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} & \begin{array}{l} s \\ m-s \end{array} \\ \begin{array}{cc} \;t\; & \;l-t\; \end{array} & \end{array},$$

则有

$$\underset{n\times m}{A}\underset{m\times l}{B}=\begin{pmatrix} A_{11}B_{11}+A_{12}B_{21} & A_{11}B_{12}+A_{12}B_{22} \\ A_{21}B_{11}+A_{22}B_{21} & A_{21}B_{12}+A_{22}B_{22} \end{pmatrix}.$$

这就告诉我们矩阵分块相乘时，可以把每一子块看成“元素”，和不分块的乘法一样进行，唯一要注意的只是乘法的顺序不能随便颠倒.

利用分块乘法的性质，容易导出下面一系列的公式：

(i) 设 $A=\begin{array}{cl}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} & \begin{array}{l} r \\ n-r \end{array} \\ \begin{array}{cc} \;r\; & \;n-r\; \end{array} & \end{array}$，且 $|A_{11}|\neq 0$，则

$$\begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix} =\begin{pmatrix} A_{11} & 0 \\ 0 & A_{22}-A_{21}A_{11}^{-1}A_{12} \end{pmatrix}. \tag{4.1}$$

这只需直接验证就可以了.对上式两边求逆(假定 A^{-1}存在)，就得

(ii) 设同(i)，A^{-1}存在时，就有

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1}=\begin{pmatrix} A_{11}^{-1}+A_{11}^{-1}A_{12}B^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}B^{-1} \\ -B^{-1}A_{21}A_{11}^{-1} & B^{-1} \end{pmatrix}$$

$$=\begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}+\begin{pmatrix} A_{11}^{-1}A_{12} \\ -I \end{pmatrix}B^{-1}(A_{21}A_{11}^{-1} \quad -I), \tag{4.2}$$

其中 $B=A_{22}-A_{21}A_{11}^{-1}A_{12}$.

类似地,设$|A_{22}|\neq 0$, A^{-1}存在,则有

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1}=\begin{pmatrix} 0 & 0 \\ 0 & A_{22}^{-1} \end{pmatrix}+\begin{pmatrix} -I \\ A_{22}^{-1}A_{21} \end{pmatrix}D^{-1}(-I \quad A_{12}A_{22}^{-1}), \tag{4.3}$$

其中 $D=A_{11}-A_{12}A_{22}^{-1}A_{21}$.

(iii) 比较(4.2)与(4.3)式中右端两表达式左上角的子块,将 A_{11}, A_{22}, A_{12}, A_{21}分别写 F, G, H, K,于是有:

如 F^{-1}, G^{-1}存在,则当$(F-HG^{-1}K)^{-1}$存在时,就有

$$(F-HG^{-1}K)^{-1}=F^{-1}+F^{-1}H(G-KF^{-1}H)^{-1}KF^{-1}. \tag{4.4}$$

如 $G=I$,就有

$$(F-HK)^{-1}=F^{-1}+F^{-1}H(I-KF^{-1}H)^{-1}KF^{-1}; \tag{4.5}$$

如 $H=u$, $K=v'$,则有

$$(F-uv')^{-1}=F^{-1}+\frac{1}{1+v'F^{-1}u}(F^{-1}uv'F^{-1}). \tag{4.6}$$

利用这些分块求逆的公式,可以得到一些行列式的公式,对(i)中等式两边取行列式,就得

(iv) $\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix}=|A_{11}||A_{22}-A_{21}A_{11}^{-1}A_{12}| \quad (|A_{11}|\neq 0)$,自然也有

$$\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix}=|A_{22}||A_{11}-A_{12}A_{22}^{-1}A_{21}| \quad (|A_{22}|\neq 0),$$

因此,当$|A_{11}|\neq 0$, $|A_{22}|\neq 0$ 时,就有

$$|A_{11}||A_{22}-A_{21}A_{11}^{-1}A_{12}|=|A_{22}||A_{11}-A_{12}A_{22}^{-1}A_{21}|$$

$$=\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix}. \tag{4.7}$$

取 $A=\begin{pmatrix} I_n & -M \\ N & I_m \end{pmatrix}$,就有

$$\begin{vmatrix} I_n & -M \\ N & I_m \end{vmatrix}=|I_m+NM|=|I_n+MN|, \tag{4.8}$$

当 $n=1$ 时,记 $M=x'$, $N=y$,上式即

$$\begin{vmatrix} 1 & -x' \\ y & I_m \end{vmatrix}=|I_m+yx'|=(1+x'y). \tag{4.9}$$

类似地可以证得当 F^{-1}存在时,

$$-x'F^{-1}y=1-|F|^{-1}|F+yx'|. \tag{4.10}$$

4.3 初等变换下的标准型 考虑对矩阵 $\underset{n\times m}{A}$ 左乘一个非奇异阵 P,右乘一个非奇异阵 Q,即将 $A\to\underset{n\times n}{P}\underset{n\times m}{A}\underset{m\times m}{Q}$($P^{-1},Q^{-1}$都存在),讨论在这种变换下 A 可以化简到什么形式.

将 A 的列进行置换,可以把

$$A\to(\underset{r}{A_1}\quad\underset{m-r}{A_2})\text{使 }\mathrm{rk}(A_1)=r=\mathrm{rk}(A),$$

即存在 Q_1,使 $AQ_1=(A_1\quad A_2)$,于是 $A_2=A_1B_1$(因为 A_2 中的每一列均可被 A_1 中的列线性表示),取 $Q_2=\begin{pmatrix}I_r & -B_1\\ O & I_{m-r}\end{pmatrix}$,则 $AQ_1Q_2=(A_1\quad A_2)\begin{pmatrix}I & -B_1\\ 0 & I_{m-r}\end{pmatrix}=(A_1\quad 0)$.类似地,对$(A_1\quad 0)$进行“行置换”,即存在 P_1 使

$$P_1(A_1\quad 0)=\begin{array}{c c}\begin{pmatrix}A_{11} & 0\\ A_{21} & 0\end{pmatrix} & \begin{matrix}r\\ n-r\end{matrix}\\ \begin{matrix}\ \ r & \qquad m-r\end{matrix} & \end{array},\text{且 }\mathrm{rk}(A_{11})=r=\mathrm{rk}\,A_1,\text{因此,存在 }B_2\text{ 使}$$

$$A_{21}=B_2A_{11},$$

取 $P_2=\begin{pmatrix}I_r & O\\ -B_2 & I_{n-r}\end{pmatrix}$,于是

$$P_2P_1(A_1\quad 0)=\begin{pmatrix}I_r & 0\\ -B_2 & I_{n-r}\end{pmatrix}\begin{pmatrix}A_{11} & 0\\ A_{21} & 0\end{pmatrix}=\begin{pmatrix}A_{11} & 0\\ 0 & 0\end{pmatrix}.$$

注意到 P_1,P_2,Q_1,Q_2 实际上都是一些初等变换的乘积,因此取 $P=P_2P_1$, $Q=Q_1Q_2$,则有

$$PAQ=\begin{array}{c c}\begin{pmatrix}A_{11} & 0\\ 0 & 0\end{pmatrix} & \begin{matrix}r\\ n-r\end{matrix}\\ \begin{matrix}\ \ r & \qquad m-r\end{matrix} & \end{array},\ |A_{11}|\neq 0.$$

回忆一下,通常用消去法解线性方程组的过程,就可以理解:对 $n\times n$ 的方阵 A, $|A|\neq 0$,可以用一系列的初等变换把它变为单位阵 I_n.对 A_{11}应用这一结论,就知道,对任给的 $\underset{n\times m}{A}$,存在 P 和 Q, $|P|\neq 0$, $|Q|\neq 0$,使

$$PAQ=\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix},\text{其中 }r=\mathrm{rk}(A). \tag{4.11}$$

$\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix}$这一类型的矩阵就是在初等变换下的标准型.注意到 P,Q 有逆,上式即

$$A=P^{-1}\begin{pmatrix}I_r & 0\\ 0 & 0\end{pmatrix}Q^{-1},$$

也即 A 是由标准型经过一系列的初等变换变来的.

§5. 特征根及特征向量

给定一个 $n\times n$ 的方阵 A，$|\lambda I-A|$是 λ 的 n 次多项式，称它为 A 的特征多项式，$|\lambda I-A|=0$ 称为 A 的特征方程. 特征方程的解，也就是特征多项式的根，称为 A 的特征根. 设 λ 是 A 的特征根，此时 $|\lambda I-A|=0$，因而方程

$$(\lambda I-A)x=0$$

一定有非零解，λ 所相应的非 0 解向量称为 λ 相应的特征向量，有时也称它为矩阵 A 的特征向量. 要注意的是特征多项式$|\lambda I-A|$虽然是实系数的多项式，但它的根不一定是实数，这样涉及到矩阵的特征根时，有时会越出我们一开始所规定的实数域的范围.

5.1 迹　对任一给定的 $n\times n$ 方阵，A 的迹 $\mathrm{tr}(A)$是 A 的全部特征根之和. 对特征多项式利用根与系数的关系式，就知道 $\mathrm{tr}(A)=\sum_{i=1}^{n}a_{ii}$. 容易验证 $\mathrm{tr}(A)$具有下列性质：

(i) $\mathrm{tr}(A+B)=\mathrm{tr}(A)+\mathrm{tr}(B)$；

(ii) $\mathrm{tr}(cA)=c\,\mathrm{tr}(A)$.

如果 A,B 是两个大小相同的 $n\times m$ 的矩阵，于是 AB'与 $A'B$ 分别为 $n\times n$ 与 $m\times m$ 的方阵.

令

$$|\lambda I_m-A'B|=|I_n||\lambda I_m-A'B|=\begin{vmatrix}\lambda I_m & A'\\ B & I_n\end{vmatrix}$$

$$=|\lambda I_m|\left|I_n-\frac{1}{\lambda}BA'\right|=\lambda^{m-n}|\lambda I_n-BA'|,\lambda\neq 0,$$

由此可见 AB'与 $A'B$ 的特征多项式只差 λ^{m-n}这个因式，因而它们的非 0 特征根全部相同，于是就得一重要的结论：

当 AB 与 BA 这两个乘积有意义时，它们的非 0 特征根全部相同.

要注意的是：如果 λ_0 是 AB 的非 0 特征根，λ_0 可以是多项式$|\lambda I-AB|$的 l 重根，上述结论告诉我们 λ_0 一定也是多项式$|\lambda I-BA|$的 l 重根. 由此立即得到：当 AB 和 BA 均为方阵(但大小不要相同)时，就有

$$\mathrm{tr}(AB)=\mathrm{tr}(BA).\tag{5.1}$$

这一等式是常常用到的，例如 $\mathrm{tr}(AA')=\mathrm{tr}(A'A)$，$\mathrm{tr}(ab')=\mathrm{tr}(b'a)=b'a$(因为向量 a,b 的内积 $b'a$ 是一个数).

5.2 海密尔顿-凯莱定理　设 $\underset{n\times n}{A}$ 的特征多项式是 $\varphi_A(\lambda)$，则 $\varphi_A(\lambda)$在复域中可分解为 n 个一次因式的乘积，即

$$\varphi_A(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n),$$

这些 λ_i 就是 A 的特征根. 形式地,可以定义

$$\varphi_A(X)=(X-\lambda_1 I)(X-\lambda_2 I)\cdots(X-\lambda_n I).$$

设 A_{ij}是 A 的元素 a_{ij}的代数余子式,记

$$\tilde{A}=\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix},$$

$\tilde{A}$ 称为 A 的伴随矩阵,于是有

$$A\tilde{A}=\begin{pmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & |A| \end{pmatrix}=|A|\cdot I_n.$$

记 $\lambda I-A$ 的代数余子式构成的伴随矩阵为 B,于是有$(\lambda I-A)B=|\lambda I-A|\cdot I_n$. 显然,$B$ 是 λ 的 $n-1$ 次多项式形成的矩阵,因此可写成 $B_0+\lambda B_1+\cdots+\lambda^{n-1}B_{n-1}$,代入上式,得

$$\begin{aligned}(\lambda I-A)(B_0+\lambda B_1+\cdots+\lambda^{n-1}B_{n-1})&=|\lambda I-A|\cdot I_n\\ &=\varphi_A(\lambda)\cdot I_n=\alpha_0 I+\alpha_1\lambda I+\cdots+\alpha_{n-1}\lambda^{n-1}I+\lambda^n I.\end{aligned}$$

比较 λ 的系数矩阵,得等式:

$$\begin{aligned} -AB_0 &= \alpha_0 I,\\ -AB_1+B_0 &= \alpha_1 I,\\ &\vdots\\ -AB_{n-1}+B_{n-2} &= \alpha_{n-1}I,\\ B_{n-1} &= I,\end{aligned}$$

上列各式依次左乘 $I, A, A^2, \cdots, A^{n-1}$,然后相加,得 $0=\varphi_A(A)$,这就是要证明的海密尔顿-凯莱定理. 定理断言如果 $\varphi_A(\lambda)$是 A 的特征多项式,则一定有 $\varphi_A(A)=0$.

于是立刻可知 $\varphi_A(A)x=0$,对一切 x 成立. 这也就告诉我们,如果 $\lambda_1,\cdots,\lambda_n$ 是 A 的特征根,则

$$(A-\lambda_1 I)(A-\lambda_2 I)\cdots(A-\lambda_n I)x=\varphi_A(A)x=0,$$

取 $u=(A-\lambda_2 I)(A-\lambda_3 I)\cdots(A-\lambda_n I)x$,则$(A-\lambda_1 I)u=0$. 很明显 $u\in\mathscr{L}(x, Ax, A^2x,\cdots)$. 这就告诉我们,对任一 $x\neq 0$,一定有 A 的一个特征向量属于 $\mathscr{L}(x, Ax,\cdots)$.

5.3 谱分解 设 $\underset{n\times n}{A}$ 的特征根是 λ,相应的特征向量是 u,于是

$$Au = \lambda u.$$

然而，A'的特征多项式与 A 的特征多项式是相同的，因为$|\lambda I - A| = |\lambda I - A'|$，于是 λ 也是 A'的特征根，因此，存在 v 使

$$A'v = \lambda v.$$

这就告诉我们，对 A 的每一个特征根 λ_i，存在 u_i 及 v_i 使 $Au_i = \lambda_i u_i$，$A'v_i = \lambda_i v_i$.

A 的特征根如果全不相同，设为 $\lambda_1, \cdots, \lambda_n$，于是就有 $2n$ 个向量 $u_i, v_i, i=1, 2, \cdots, n$，使

$$Au_i = \lambda_i u_i, A'v_i = \lambda v_i, i = 1, 2, \cdots, n.$$

记 $U = (u_1 \cdots u_n)$，$V = (v_1 \cdots v_n)$，容易看出，

$$AU = U\Lambda, A'V = V\Lambda,$$

其中 $A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$. 现在证明 U^{-1}，V^{-1}是存在的，也即要证 $u_1, \cdots, u_n$ 是线性无关的，$v_1, \cdots, v_n$ 是线性无关的. 若不然，存在 $c_1, \cdots, c_n$ 不全为 0 使

$$\sum_{i=1}^{n} c_i u_i = 0,$$

于是 $0 = A0 = A\left(\sum_{i=1}^{n} c_i u_i\right) = \sum_{i=1}^{n} c_i \lambda_i u_i$. 由 $\sum_{i=1}^{n} c_i u_i = 0$ 及 $\sum_{i=1}^{n} c_i \lambda_i u_i = 0$，可以消去某一个 u_i，无妨设消去 u_1，于是有 d_i 不全为 0 使 $\sum_{i=2}^{n} d_i u_i = 0$，再重复刚才的做法，……，最后逐一消去，得 $u_n = 0$，这是不可能的，因此 U 一定有逆. 同理 V 也有逆. 又因为

$$\lambda_i v'_i u_j = v'_i A u_j = \lambda_j v'_i u_j,$$

于是，当 $i \neq j$ 时 $v'_i u_j = 0$ 对一切 $i, j = 1, 2, \cdots, n$ 成立，也即

$$V'U = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} \triangleq D.$$

易见 D^{-1}存在，而且 $D^{-1} = U^{-1}(V')^{-1}$，即 $U^{-1} = D^{-1}V'$. 利用 $AU = U\Lambda$，U^{-1}存在，且 $U^{-1} = D^{-1}V'$，就有

$$A = U\Lambda U^{-1} = U\Lambda D^{-1}V' = U(\Lambda D^{-1})V' = \sum_{i=1}^{n} \frac{\lambda_i}{d_i} u_i v'_i.$$

设 u_i 是 A 的特征向量，则 cu_i 一定也还是 A 的特征向量，适当选取 u_i 及 v_i 使 $v'_i u_i = 1, i = 1, 2, \cdots, n$，于是 $d_i \equiv 1$，因此

$$A = \sum_{i=1}^{n} \lambda_i u_i v'_i. \tag{5.2}$$

这就是矩阵 A 的谱分解,特征根 $\lambda_1,\cdots,\lambda_n$ 也称为 A 的谱.

如果 A 的特征根有重根,例如 λ_i 是 A 的 l_i 重根,如果相应于 λ_i 有 l_i 个线性无关的特征向量,那么类似的讨论仍可进行.实际上,由于 A 的特征根的重数与它线性无关的特征向量的个数不一定是相同的,因此谱分解式在一般情况下并不成立.很明显,矩阵

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

的三个特征根是 1,1,1,然而 1 相应的特征向量 $x=(x_1,x_2,x_3)'$ 是

$$\begin{pmatrix}1&0&0\\1&1&0\\0&1&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$$

$$\Leftrightarrow x_1=x_1,\ x_1+x_2=x_2,\ x_2+x_3=x_3$$

$$\Leftrightarrow x_1=x_2=0,\ x_3\ \text{任意},$$

它只有一个非 0 特征向量.

5.4 特征根与行列式的关系 从矩阵 $\underset{n\times n}{A}$ 的特征多项式 $|\lambda I-A|$ 可以看出,它的常数项就是行列式 $|A|$ 乘以 $(-1)^n$;另一方面,从根与系数的关系又知道,常数项是全部特征根的乘积乘以 $(-1)^n$.设 A 的全部特征根为 $\lambda_1,\cdots,\lambda_n$,于是有

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

因此 $|A|=\prod_{i=1}^n\lambda_i$,于是就可以推出

$$A\ \text{非奇异}\Leftrightarrow A\ \text{的特征根均不为 0},$$

$$A\ \text{奇异}\Leftrightarrow A\ \text{至少有一个特征根是 0}.$$

5.5 等幂阵 如果方阵 $\underset{n\times n}{A}$ 具有性质 $A^2=A$,则称 A 是等幂阵.由于等幂阵和投影有密切的关系,这里先概述一些等幂阵的性质.

(i) 如果 $A^2=A$,则 A 的特征根非 0 即 1.这是因为当 λ 是 A 的特征根时,就有 $u\neq0$ 使 $Au=\lambda u$.于是 $A^2u=A(\lambda u)=\lambda^2u$,也即 $Au=\lambda^2u$,因此得 $\lambda^2=\lambda$,也即 λ 非 0 即 1.

(ii) 如果 $A^2=A$,则 $\mathrm{rk}(A)=\mathrm{tr}(A)$.利用初等变换可以将矩阵 A 写成 $A=P\begin{pmatrix}I_r&0\\0&0\end{pmatrix}Q$,$|P|\neq0$,$|Q|\neq0$,$r=\mathrm{rk}(A)$,由 $A^2=A$,得

$$P\begin{pmatrix}I_r&0\\0&0\end{pmatrix}QP\begin{pmatrix}I_r&0\\0&0\end{pmatrix}Q=P\begin{pmatrix}I_r&0\\0&0\end{pmatrix}Q,$$

因此

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} QP \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

把 Q,P 分块,记 $P=(\underset{r}{P_1}\quad \underset{n-r}{P_2})$, $Q=\begin{array}{c@{}l} \begin{pmatrix} Q_1' \\ Q_2' \end{pmatrix} & \begin{array}{l} r \\ n-r \end{array} \end{array}$,于是

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} Q_1' P_1 & Q_1' P_2 \\ Q_2' P_1 & Q_2' P_2 \end{pmatrix} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

也即 $Q_1' P_1 = I_r$.记 $P_1=(p_1\ \ p_2 \cdots p_r)$, $Q_1=(q_1\ \ q_2\cdots q_r)$,则有 $p_i'q_j = q_j'p_i = \delta_{ij}$, i, $j=1,\cdots,r$,而

$$A = P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q = (P_1 \quad P_2)\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} Q_1' \\ Q_2' \end{pmatrix} = P_1 Q_1',$$

即

$$A = \sum_{i=1}^{r} p_i q_i',$$

而 $\operatorname{tr} A = \sum_{i=1}^{r} \operatorname{tr} p_i q_i' = \sum_{i=1} \operatorname{tr} q_i' p_i = r = \operatorname{rk}(A)$.这也就顺便得出了

(iii) $A^2 = A \Leftrightarrow A = \sum_{i=1}^{r} p_i q_i'$, $p_i' q_j = \delta_{ij}$, $i,j=1,2,\cdots,r$.容易看出 $A^2 = A$ 时,$(I-A)^2 = I-A$,于是有

$$\begin{aligned} n &= \operatorname{rk} I_n = \operatorname{tr}(I_n - A + A) = \operatorname{tr}(I-A) + \operatorname{tr} A \\ &= \operatorname{rk}(I-A) + \operatorname{rk}(A), \end{aligned}$$

而$\mathscr{L}(A')$的维数是 $\operatorname{rk}(A)$,$\mathscr{L}(I-A)$的维数是 $\operatorname{rk}(I-A)$.设 $x\in\mathscr{L}(A')$, $y\in\mathscr{L}(I-A)$,则 $x=A'u$, $y=(I-A)v$,于是

$$x'y = u'A(I-A)v = 0,$$

即$\mathscr{L}(A')$中的向量与$\mathscr{L}(I-A)$中的向量正交,即 R_n 是$\mathscr{L}(A')$与$\mathscr{L}(I-A)$的正交和,因此有

(iv) $R_n = \mathscr{L}(A') \dotplus \mathscr{L}(I-A) = \mathscr{L}(A) \dotplus \mathscr{L}(I-A') \Leftrightarrow A^2 = A$.

§6. 对　称　阵

方阵 A 称为对称的,如果 $A'=A$.对每一个对称阵 A,任给一个向量 x, $x'Ax$ 就是 x 的一个齐次二次函数,它称为 A 所相应的二次型.如果 A 的二次型 $x'Ax$ 恒不取负值,即 $x'Ax\geqslant 0$ 对一切 x 成立,则称 A 是非负定阵.如果 A 是非负定的,且 $x'Ax=0$ 的充要条件是 $x=0$,则称 A 是正定的.非负定阵和正定阵都是对称阵,这一节着重介绍有关特征根,特征向量的一些结果.

6.1 对称阵的谱分解 首先可以证明对称阵的特征根均取实值. 如果 λ, μ, x, y 分别为实数及实数向量, $A' = A$, 且

$$A(x + iy) = (\lambda + i\mu)(x + iy), \text{其中 } i \text{ 为虚数}.$$

比较等式两边的实部及虚部, 得

$$Ax = \lambda x - \mu y,\ Ay = \mu x + \lambda y,$$

所以
$$\lambda y'x - \mu y'y = y'Ax = x'Ay = \mu x'x + \lambda x'y,$$
因此
$$\mu(x'x + y'y) = 0,$$
即
$$\mu = 0.$$

这就证明了特征根一定是实的, 同时特征向量也可取为实值向量.

其次 $\underset{n\times n}{A}' = A$ 时, 存在 A 的 n 个特征向量 $\gamma_1, \cdots, \gamma_n$ 使 $\gamma_i'\gamma_j = \delta_{ij}$, $i, j = 1, 2, \cdots, n$. 设 A 的 k 个单位正交特征向量为 $\gamma_1, \cdots, \gamma_k$(显然, 对 A 至少可以找到一个单位长度的特征向量, 因此 $k \geqslant 1$), 相应的特征根是 $\lambda_1, \cdots, \lambda_k$($\lambda_1, \cdots, \lambda_k$ 可以有相同的). 于是当 $k < n$ 时, 总有 $x \neq 0$, 使 x 与 $\mathscr{L}(\gamma_1, \cdots, \gamma_k)$ 中的向量都正交, 而 $(Ax)'\gamma_i = x'A\gamma_i = \lambda_i x'\gamma_i = 0$, $i = 1, 2, \cdots, k$, 即 $x, Ax, \cdots$, 均与 $\mathscr{L}(\gamma_1, \cdots, \gamma_k)$ 正交. 由 §5 的 5.2 知道 $\mathscr{L}(x, Ax, \cdots)$ 中一定有 A 的特征向量, 因此就找到了与 $\gamma_1, \cdots, \gamma_k$ 都正交的 A 的单位特征向量 γ_{k+1}. 用这个方法直到 $k = n$ 为止, 就找到了 A 的 n 个特征向量 $\gamma_1, \cdots, \gamma_n$, 它们是 R_n 中的一组标准正交基, 即 $\delta_{ij} = \gamma_i'\gamma_j$, $i, j = 1, \cdots, n$, 记

$$\Gamma = (\gamma_1 \cdots \gamma_n),$$

则有 $A\Gamma = \Gamma\Lambda$, 其中 $\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, 且 $\Gamma'\Gamma = I_n$. 利用逆阵的唯一性, 就知道 $\Gamma' = \Gamma^{-1}$, $\Gamma\Gamma' = I_n$. 于是

$$A = \Gamma\Lambda\Gamma^{-1} = \Gamma\Lambda\Gamma' = (\gamma_1 \cdots \gamma_n)\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}\begin{pmatrix} \gamma_1' \\ \vdots \\ \gamma_n' \end{pmatrix}$$

$$= \sum_{i=1}^{n} \lambda_i \gamma_i \gamma_i'.$$

这就证得了 A 的谱分解式.

A 的谱分解也可以用另一种方式来叙述. 设 Γ 是 $n \times n$ 的阵, 如果 $\Gamma'\Gamma = \Gamma\Gamma' = I_n$, 则称 Γ 是正交阵. A 的谱分解也可以叙述为:

对任给的一个对称阵 $\underset{n\times n}{A}$, 一定存在一个正交阵 Γ, 使

$$\Gamma' A \Gamma = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \tag{6.1}$$

其中 $\lambda_1,\cdots,\lambda_n$ 是 A 的特征根，$\Gamma=(\gamma_1\cdots\gamma_n)$是相应的特征向量组成的正交阵.这也称为对称阵的对角化定理.

对给定的对称阵 A，任给一个正交阵 Γ，则 $\Gamma' A\Gamma$ 仍然是对称阵.上述结果告诉我们，对 $A\to\Gamma' A\Gamma$ 这种变换，总可以使 A 变成对角形的阵，因此对角阵称为对称矩阵在正交变换下的标准型.

6.2 同时对角化　设 A,B 都是 $n\times n$ 的对称阵.对于 A，存在正交阵 Γ_1 使 $\Gamma_1' A\Gamma_1$ 是对角阵；对于 B，存在正交阵 Γ_2 使 $\Gamma_2' B\Gamma_2$ 也是对角阵.一般说来 Γ_1 与 Γ_2 是不同的.下面证明，$\Gamma_1=\Gamma_2$ 的充要条件是 $AB=BA$.这一结果通常称为同时对角化定理.

先证必要性.若 $\Gamma_1=\Gamma_2$，则 $AB=BA$.此时

$$\Gamma_1' A\Gamma_1=\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix},\quad \Gamma_2' B\Gamma_2=\begin{pmatrix}\mu_1 & & 0\\ & \ddots & \\ 0 & & \mu_n\end{pmatrix},$$

由于 $\Gamma_1=\Gamma_2$，记 $\Gamma_1=\Gamma_2=\Gamma$，则

$$A=\Gamma\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}\Gamma',\quad B=\Gamma\begin{pmatrix}\mu_1 & & 0\\ & \ddots & \\ 0 & & \mu_n\end{pmatrix}\Gamma',$$

因此

$$AB=\Gamma\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}\Gamma'\Gamma\begin{pmatrix}\mu_1 & & 0\\ & \ddots & \\ 0 & & \mu_n\end{pmatrix}\Gamma'=\Gamma\begin{pmatrix}\lambda_1\mu_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\mu_n\end{pmatrix}\Gamma'$$

$$=\Gamma\begin{pmatrix}\mu_1 & & 0\\ & \ddots & \\ 0 & & \mu_n\end{pmatrix}\Gamma'\Gamma\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}\Gamma=BA.$$

再证充分性.若 λ_1,u_1 使 $Au_1=\lambda_1 u_1$，则 $AB^k u_1=B^k Au_1=\lambda_1 B^k u_1$.可见当 u_1 是 λ_1 相应的 A 的特征向量时，$Bu_1,B^2 u_1,\cdots,B^k u_1,\cdots$都在 λ_1 相应的特征向量生成的子空间内.由 5.2 知道一定有一个 B 的特征向量属于 $\mathscr{L}(u_1,Bu_1,B^2 u_1,\cdots)$，于是就找到了 A,B 第一个公共的特征向量，记为 γ_1.考虑与 γ_1 正交的子空间，存在 u_2 是 A 的特征向量，与刚才的方法相仿，可以找到 A,B 第二个公共特征向量 γ_2，且 $\gamma_2'\gamma_1=0$.如此下去，一直找到 n 个公共的特征向量为止，这就证明了所要的结论.

因此，当 $A_1,\cdots,A_m$ 均为 $n\times n$ 的对称阵时，$A_1,\cdots,A_m$ 可以同时对角化的充要条件是 $A_iA_j=A_jA_i$ 对一切 $i,j=1,2,\cdots,m$ 成立.特别地，对称阵 A 与 $A^2,A^3,\cdots,A^{-1},A^{-2},\cdots$都是可交换的，因此 A 与 A 的多项式矩阵，或 A^{-1}的多项式矩阵都

可以同时对角化.用谱分解的形式来表示,如果 A 的特征根是 $\lambda_1,\cdots,\lambda_n$,$\gamma_1,\cdots,\gamma_n$ 是相应的特征向量(标准正交基).当 $\lambda_i=0$ 时,规定 $\lambda_i^{-1}=0$,于是有

$$\begin{cases} A=\sum_{i=1}^{n}\lambda_i\gamma_i\gamma_i'; \\ A^k=\sum_{i=1}^{n}\lambda_i^k\gamma_i\gamma_i',\ k=0,\pm1,\pm2,\cdots. \end{cases} \tag{6.2}$$

如果 $f(\lambda)$是 λ 的多项式,则有

$$\begin{cases} f(A)=\sum_{i=1}^{n}f(\lambda_i)\gamma_i\gamma_i', \\ f(A^{-1})=\sum_{i=1}^{n}f(\lambda_i^{-1})\gamma_i\gamma_i'. \end{cases} \tag{6.3}$$

6.3 特征根的极值性质 设 $A'=A$,将 A 的特征根 $\lambda_1,\cdots,\lambda_n$ 依大小顺序排列,无妨设 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$.从谱分解式($\gamma_1,\cdots,\gamma_n$ 是 A 的标准化特征向量):

$$A=\sum_{i=1}^{n}\lambda_i\gamma_i\gamma_i',$$

$$I=\sum_{i=1}^{n}\gamma_i\gamma_i',$$

就可以知道,对任给 x,$x=\sum_{i=1}^{n}a_i\gamma_i$ 就有

$$\frac{x'Ax}{x'x}=\frac{\sum_{i=1}^{n}\lambda_i a_i^2}{\sum_{i=1}^{n}a_i^2},$$

很明显,利用上面的等式就可以得到

$$\begin{cases} \sup\limits_{x\neq0}\dfrac{x'Ax}{x'x}=\sup\limits_{a\neq0}\dfrac{\sum_{i=1}^{n}\lambda_i a_i^2}{a'a}=\lambda_1, \\ \inf\limits_{x\neq0}\dfrac{x'Ax}{x'x}=\inf\limits_{a\neq0}\dfrac{\sum_{i=1}^{n}\lambda_i a_i^2}{a'a}=\lambda_n. \end{cases} \tag{6.4}$$

仿照上面的方法,不难证明下述结论:

设 $\underset{n\times n}{A'}=A$,$A$ 的谱分解式是 $\sum_{i=1}^{n}\lambda_i\gamma_i\gamma_i'$,且 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$,于是有:

(i) $$\sup_{\substack{x'\gamma_i=0\\ i=1,\cdots,k\\ x\neq0}}\frac{x'Ax}{x'x}=\lambda_{k+1},\quad \inf_{\substack{x'\gamma_i=0\\ i=1,\cdots,k\\ x\neq0}}\frac{x'Ax}{x'x}=\lambda_n; \tag{6.5}$$