常用的矩阵理论和方法

倪国熙编

上海科学技术出版社

常用的矩阵理论和方法

倪国熙　编

上海科学技术出版社出版

（上海瑞金二路 450 号）

新华书店上海发行所发行　上海印刷三厂印刷

开本 787×1092 1/32　印张 6.625　字数 145,000

1984 年 4 月第 1 版　1984 年 4 月第 1 次印刷

印数：1—20,400

统一书号：13119·1140　定价：（科五）0.76元

前　言

随着数学广泛应用于各个学科，矩阵的重要性正在与日俱增，矩阵的理论和方法亦在不断的发展之中，从而使教科书中所讲述的矩阵内容，已日益显示出不能满足实际的需要。撰写此书的目的，就是企图从某个侧面去弥补此种不足。

我们希望本书成为大学生在矩阵方面的一本课外读物，同时也为有关的高等学校教师和科技工作者提供一本有益的参考书。因而，它在取材和阐述方面，既不能类同于一般教科书，也应当有别于专门著作。本书立足于矩阵的常用理论和方法，但侧重点在于一般教科书中未予涉及或强调不够的，而在实际上却十分有用的那些部分，如分块矩阵、投影阵、广义逆矩阵、矩阵的不等式与极值以及矩阵的特殊运算等，并尽力反映较新的方法和结果。为了使本书能自成体系，便利于阅读和参考，也因为在观点和方法上的特色，我们用将近三分之一的篇幅阐述了矩阵的基础理论。相信这部分材料，对于已经学过线性代数的读者，也会有所裨益。其余各章，相互联系并不紧密。由于受篇幅的限制，若干章节未予充分展开，因而长短不一，写法也有区别。

考虑到本书的对象大多有一定的代数基础，全书的叙述力求扼要。虽然在概念和推理上尽量做到严谨，但具体的例子极少，一些有用的结论又常常列为思考题、习题或章末的问题和补充，从而使本书具有一定的难度，在阅读时需要相当的耐心，要勤于思索和动笔，否则恐怕难以有真正的收获。作者认为，要罗列矩阵方面的异常丰富的结果是不可能的，即使

在我们所涉及的较狭窄的论题下，对结论的叙述也难免挂一漏万。学习的成效，不在于记住了多少结论，而是要透彻地掌握理论和方法，不断提高分析问题与解决问题的能力。唯有如此，才能在变化无穷的实际问题面前，做到举一反三，得心应手。因此，要求读者在阅读过程中能独立地举一些正反两方面的例子，做一些较难的练习，在费力的攀登中锻炼自己的功力。

由于本书的性质，未列出详尽的参考文献。对于部分地使用的其它书的材料，也恕不予以注明。读者欲进一步钻研，可从书末所列的主要参考书中查找有关文献。

本书是在选修课讲义的基础上，融合作者在工作中的体会发展起来的，肯定有错误和不足，敬祈有关专家和读者不吝批评指正。

在酝酿和写作本书的过程中，得到武汉大学数学系张尧庭教授的大力支持。研究生丁树良协助做了不少具体工作，特在此志谢。

倪 国 熙

1982 年 2 月

目　　录

第一章　矩阵的基础理论

为了尽量不引用其它书中的结论，并且贯串我们所强调的观点和方法，将在本章中扼要叙述矩阵的基础理论，其内容相当于一般高等代数教材中所包含的矩阵部分。我们将在线性空间观点下展开矩阵的基础理论。在方法上强调了分块矩阵的运用，即把不必过细剖分的部分当作一个整体来处理。这样既避免了繁琐的表达式，又易于抓住主要特征。为了节省篇幅，只给出一些关键的或有特色的定理，并省略了许多一般化的证明，或仅仅给出证明的提要。小节后面的习题与章末的问题，大部分是正文的补充，但也有少量问题较难。证明这些结果，可作为对运用理论与方法的练习。

矩阵是一个十分常见的数学对象。将 $m\times n$ 个元素排成 m 行、n 列的一个矩形阵，就称为**矩阵**。如

$$A \triangleq \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}^{(*)},$$

其中的 a_{ij} 称为矩阵 A 的 (i,j)–元，即 A 中处于第 i（横）行、第 j（纵）列位置的元素。常简记为 $A\triangleq[a_{ij}]$。矩阵的元素可以取自抽象的代数系统，如域和环等。但在本书中，只讨论元素为实数或复数的矩阵。并且，为了叙述方便，在不加说明时，我们仅讨论实数阵。$m\times n$ 称为矩阵 A 的阶。矩阵 A 与 B 相等，意味着它们的阶相同，且各个元素对应相等。

(*) 今后，凡定义式或记号，常用符号≜表示。

由 A 导出的 $n\times m$ 阶阵

$$A^T \triangleq \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

称作 A 的**转置**。显然有 A^T 的 (i, j)-元是 a_{ji}，且有 $(A^T)^T = A$。

一个 $p\times 1$ 阶的矩阵称作 p 维**向量**，如

$$\boldsymbol{b} \triangleq \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{bmatrix} = [b_1 b_2 \cdots b_p]^T,$$

其中 b_i 称作 $\boldsymbol{b}$ 的第 i 个**分量**。

将矩阵 A 剖分为若干个较低阶矩阵（子阵）的做法是十分有用的。经过剖分的矩阵称作**分块矩阵**。如将 A 剖分为

$$A \triangleq \begin{array}{c} \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \begin{array}{l} p \\ \\ \end{array} \\ \begin{array}{c} q \qquad\qquad \end{array} \end{array},$$

其中 A_{11} 是 $p\times q$ 阶，其余子阵的阶数相应可得，如 A_{21} 是 $(m-p)\times q$ 阶。剖分成多少个子阵可按问题的需要而定。除上面的四块形式外，最常用的还有如下的列剖分和行剖分形式：

$$A \triangleq [\boldsymbol{a}_1 \mid \cdots \mid \boldsymbol{a}_n] \triangleq \begin{bmatrix} \boldsymbol{a}_{(1)}^T \\ \hdashline \vdots \\ \hdashline \boldsymbol{a}_{(m)}^T \end{bmatrix},$$

其中 $\boldsymbol{a}_j = [a_{1j} \cdots a_{mj}]^T$，$\boldsymbol{a}_{(i)} = [a_{i1} \cdots a_{in}]^T$，分别称作 A 的第 j 个**列向量**和第 i 个**行向量**。

上述标准记法，一般将不再重复赘述。在不致混淆时，我们将省略剖分记法中划分子阵的虚线。注意，矩阵一般用大写的英文字母表示，向量用小写黑体的英文字母，数则用小写的英文字母或希腊字母表示。处在同一矩阵中的，常用同名的字母来记。

§1　线性空间与线性映射

1-1　线性空间及其维数

线性空间是直观的二维与三维向量空间的自然的推广。在给出它的抽象定义和讨论它的种种性质时，始终有一个直观的形象作为背景是十分有益的。

定义 1.1　设 L 是一个非空集合。如果定义了 L 中二元运算，称为加法，又定义了数域 F 中的数对 L 中元素的运算，称为数乘，而加法和数乘又满足以下规律：

(1) 结合律　$(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$;

(2) 交换律　$\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$;

(3) 有零元　存在 $\mathbf{0}\in L$，使 $\mathbf{0}+\boldsymbol{a}=\boldsymbol{a}$;

(4) 有负元　$\forall \boldsymbol{a}\in L$(*)，有 $-\boldsymbol{a}$ 使得 $\boldsymbol{a}+(-\boldsymbol{a})=\mathbf{0}$;

(5) 单位数乘不变律　$1\cdot\boldsymbol{a}=\boldsymbol{a}$;

(6) 数乘结合律　$\lambda\cdot(\mu\cdot\boldsymbol{a})=(\lambda\mu)\cdot\boldsymbol{a}$;

(7) 分配律　$\lambda\cdot(\boldsymbol{a}+\boldsymbol{b})=\lambda\cdot\boldsymbol{a}+\lambda\cdot\boldsymbol{b}$,

$$(\lambda+\mu)\cdot\boldsymbol{a}=\lambda\cdot\boldsymbol{a}+\mu\cdot\boldsymbol{a};$$

则称 L 是数域 F 上的线性空间。在不致混淆时就简称线性空间。

定义中提到的数域，实质上可以是抽象的域。但与前面

(*) 符号 ∀ 念作“对一切”。

相仿，若无特别声明，我们将仅限于数域，尤其是实数域。

线性空间中的元素，通常称作**向量**。这里的向量是在抽象意义下来理解的。它与前面提到的p维向量（$p\times 1$阶矩阵），既有密切联系又有一定区别。这将在以后逐步说明。

很明显，向量的减法可规定为

$$\boldsymbol{a}-\boldsymbol{b}=\boldsymbol{a}+(-\boldsymbol{b})。$$

关于线性空间中运算（如加、减、数乘等）的一些初步性质，读者不难自行导出，我们将随意引用。

定义 1.2 设L是线性空间，$\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r\in L$。如果存在不全为零的一组数$\lambda_1, \cdots, \lambda_r$，满足

$$\sum_{i=1}^{r}\lambda_i\boldsymbol{a}_i\triangleq\lambda_1\boldsymbol{a}_1+\cdots+\lambda_r\boldsymbol{a}_r=\boldsymbol{0},$$

则称$\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r$**线性相关**或简称**相关**。否则，称$\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r$**线性无关**，或简称**无关**。

显然，线性无关的定义等价于：

$$\sum_{i=1}^{r}\lambda_i\boldsymbol{a}_i=\boldsymbol{0}\Rightarrow^{(*)}\lambda_1=\cdots=\lambda_r=0。$$

如果向量$\boldsymbol{b}=\sum_{i=1}^{r}\lambda_i\boldsymbol{a}_i$，则称$\boldsymbol{b}$可被$\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r\}$**线性表出**。当一组向量$\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_s\}$中的任一个可被$\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r\}$线性表出，称$\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_s\}$可被$\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r\}$**表出**。可相互表出的两组向量被认为是**等价**的。

这些概念是线性空间理论的出发点。下面的定理在建立线性空间的特征——维数概念时，是至关重要的。

定理 1.1 设L为线性空间，$\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r\}$、$\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_s\}\subset L$。如果$\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r\}$无关，且可被$\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_s\}$表出，则有$r\leqslant s$。且有$\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_s\}$的$r$个向量存在，不妨设为

(*) 符号⇒表示“可推出”。⇔表示可以互推；等价；充要条件。

$\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_r\}$，当用 $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r\}$ 代换它时，所得 $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r, \boldsymbol{b}_{r+1}, \cdots, \boldsymbol{b}_s\}$ 可将 $\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_s\}$ 表出。(称为代换定理)

证明可对 r 用数学归纳法得出。▌

定义 1.3 设 S 是线性空间 L 的子集，如果 S 的任意有限子集都线性无关，且 L 的任何向量均可被 S 表出，则称 S 是 L 的基。

定理 1.2 如果线性空间 L 的基 S 恰含 n 个向量，则 L 的任何基都恰含 n 个向量。

称有上述性质的线性空间为**有限维线性空间**，n 为它的**维数**，记作 $n=\dim L$。

证明易于从代换定理推出。▌

本书仅讨论有限维线性空间 L，如果需要写出它的维数，就记为 L^n。

我们约定，由单个零向量所构成的线性空间维数为零。

例 1.1 设 $R^{m\times n}$ 是实数域上 $m\times n$ 阶矩阵的全体。任给 $A=[a_{ij}]$、$B=[b_{ij}]\in R^{m\times n}$，定义

$$A+B=[a_{ij}+b_{ij}],$$

$$\mu A=[\mu a_{ij}],$$

即 $A+B$ 以 $a_{ij}+b_{ij}$ 为 (i, j)-元，μA 以 μa_{ij} 为 (i, j)-元。由此给出的 $R^{m\times n}$ 的加法和数乘，满足定义 1.1 中的(1)～(7)，因此 $R^{m\times n}$ 是线性空间。以后，我们总认为矩阵的和与数乘已按此定义。

设 E_{ij} 是 $R^{m\times n}$ 中除 (i, j)-元为 1 外，其余元素皆为 0 的矩阵，则 $S=\{E_{ij}: i=1, \cdots, m; j=1, \cdots, n\}$ 是 $R^{m\times n}$ 的基(思考题)。所以，$R^{m\times n}$ 是 $m\cdot n$ 维线性空间。

作为特例，p 维向量的全体 R^p 是 p 维线性空间。

定义 1.4 设 $S=\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$ 是 n 维线性空间 L 的基，

将 S 中向量排好次序，记为 $(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n)$，称作 L 的**有序基**。任给 $\boldsymbol{x} \in L$，必有唯一的一组数 $x_1, \cdots, x_n$，满足

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i。 \text{（思考题）}$$

称 $x_1, \cdots, x_n$ 为 $\boldsymbol{x}$ 在基 $(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n)$ 下的坐标，而 n 维向量 $[x_1, \cdots, x_n]^T$ 称作 $\boldsymbol{x}$ 在基 $(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n)$ 下的**坐标表示**。

由此可见：在确定了 n 维线性空间 L^n 的有序基之后，它的向量是和 n 维向量一一对应的。并且，这种对应保持了 L^n 与 R^n 的运算关系不变，即当 $\boldsymbol{x}$、$\boldsymbol{y} \in L^n$ 的坐标分别为 $x_1, \cdots, x_n$ 与 $y_1, \cdots, y_n$ 时，$\boldsymbol{x}+\boldsymbol{y}$ 所对应的 R^n 中的向量恰为 $[x_1+y_1, \cdots, x_n+y_n]^T$。而且数乘运算也保持不变。这类一一对应关系，在代数中称为同构。单纯从线性空间的观点来看，两个同构的线性空间被认为是没有区别的。由于实数域上的任一 n 维线性空间都同构于 R^n，故维数 n 成为线性空间的特征。

应当注意：L^n 与 R^n 的一一对应关系是依赖于有序基的选取的，或者说 L^n 中抽象向量的坐标表示是与有序基相关联的。当有序基变换时，向量的坐标将如何变换，显然是值得研究的问题。答案将在本节 1-3 中给出。

定义 1.5 设 $S \subset L^n$。如果 S 对加法和数乘封闭（即 S 中任两向量的和皆仍在 S 中，S 中任一向量与任一数的乘积亦仍在 S 中），则称 S 是 L^n 的**子空间**。显然，子空间本身是线性空间。

注意，S 是零维子空间，当且仅当 $S=\{0\}$，即单个零向量所构成的子空间。

例 1.2 T 所张成的子空间。

设 T 是 L^n 的子集。令

$$R(T)=\left\{x=\sum_{i=1}^{n}\lambda_i y_i:\ n\text{ 为任意正整数},\ y_i\in T,\ \lambda_i\text{ 是数},\ i=1,\cdots,n\right\},$$

则可验证 $R(T)$ 是 L^n 的子空间（思考题）。称 $R(T)$ 为 T 所**张成的子空间**。

例 1.3 和空间与交空间。

设 T、S 都是 L^n 的子空间。令

$$S+T=\{\boldsymbol{x}+\boldsymbol{y}:\ \boldsymbol{x}\in S,\ \boldsymbol{y}\in T\},$$

$$S\cap T=\{\boldsymbol{x}:\ \boldsymbol{x}\in S\text{ 且 }\boldsymbol{x}\in T\},$$

易证此两集合均为 L^n 的子空间（思考题），分别称作 S 与 T 的**和空间**与**交空间**。要注意空间的和与集合的并的区别。

当 $S\cap T=\{\mathbf{0}\}$，称 S 与 T 的和空间为**直和**，记作 $S\oplus T$。

习　题

1.1　设 $\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r\}$ 是 L^n 的无关集，$r<n$，则可在其中添加 $n-r$ 个向量，使 $\{\boldsymbol{a}_1,\cdots,\boldsymbol{a}_r,\boldsymbol{a}_{r+1},\cdots,\boldsymbol{a}_n\}$ 为 L^n 的基。

1.2　设 S、T 是 L^n 的子空间，且 $S\subset T$，则有

$\dim S\leqslant\dim T$。且 $\dim S=\dim T\Rightarrow S=T$。

1.3　证明

$\dim(S+T)=\dim S+\dim T-\dim(S\cap T)$。

（提示：利用习题1.1，从 $S\cap T$ 的基出发。）

1.4　$S+T$ 是直和，等价于下列条件之一：

1°　$\forall\boldsymbol{x}\in S+T$，$\boldsymbol{x}$ 可分解为 $\boldsymbol{x}=\boldsymbol{y}+\boldsymbol{z}$，$\boldsymbol{y}\in S$，$\boldsymbol{z}\in T$，且分解是唯一的；

2°　$\exists$（读作存在）$\boldsymbol{x}\in S+T$，$\boldsymbol{x}$ 可分解为 $\boldsymbol{x}=\boldsymbol{y}+\boldsymbol{z}$，$\boldsymbol{y}\in S$，$\boldsymbol{z}\in T$，且分解式是唯一的；

3° 零向量的如上分解唯一，即 $\mathbf{0}=\mathbf{0}+\mathbf{0}$;

4° $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r\}$ 与 $\{\boldsymbol{b}_1, \cdots, \boldsymbol{b}_s\}$ 分别是 S 与 T 的任两个无关集，则 $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r, \boldsymbol{b}_1, \cdots, \boldsymbol{b}_s\}$ 无关。

1-2 内积空间与正交化

直观的三维空间中，向量的长度和夹角的概念，不难推广到 n 维线性空间。为此，需要引进内积概念。

定义 1.6 设 L 是实数域 R 上的线性空间，如果定义了映射 $(\cdot, \cdot)$: $L\times L\to R$, 满足

(1) $(\boldsymbol{x}, \boldsymbol{y})=(\boldsymbol{y}, \boldsymbol{x})$,

(2) $(\alpha\boldsymbol{x}, \boldsymbol{y})=\alpha(\boldsymbol{x}, \boldsymbol{y})$,

(3) $(\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{z})=(\boldsymbol{x}, \boldsymbol{z})+(\boldsymbol{y}, \boldsymbol{z})$,

(4) $(\boldsymbol{x}, \boldsymbol{x})\geqslant 0$, 且仅当 $\boldsymbol{x}=\mathbf{0}$ 时才有等号成立，

则称此二元运算为内积。

当数域是复数域 U 时，只需将上面定义中的 R 换为 U (1) 换为

(1)′ $(\boldsymbol{x}, \boldsymbol{y})=\overline{(\boldsymbol{y}, \boldsymbol{x})}$,

同样有内积定义。

定义了内积的线性空间就称为**内积空间**。

例 1.4 标准内积

对于实数域上的 n 维向量空间 R^n, 令 $(\boldsymbol{x}, \boldsymbol{y})=\boldsymbol{y}^T\boldsymbol{x}$, 易证如此定义的二元运算为一内积。对于复数域上的 n 维向量空间 U^n, 令 $(\boldsymbol{x}, \boldsymbol{y})=\boldsymbol{y}^*\boldsymbol{x}$(这里 A^* 是矩阵 A 的**共轭转置**，即将 A 转置后再将其元素换为它的共轭复数)，亦给出一内积。这里给出的内积称为**标准内积**。上述空间在定义了标准内积后，分别称为 n 维**欧氏空间**和 n 维**酉空间**。

在内积空间中，可以自然地定义向量 $\boldsymbol{x}$ 的**范数**为：

$$\|\boldsymbol{x}\|=(\boldsymbol{x}, \boldsymbol{x})^{1/2},$$

$\|\cdot\|$ 满足三条性质：(1) $\|\boldsymbol{x}\|\geqslant 0$，且仅当 $\boldsymbol{x}=\mathbf{0}$ 时才有等号成立；(2) $\|\alpha\boldsymbol{x}\|=|\alpha|\,\|\boldsymbol{x}\|$；(3) $\|\boldsymbol{x}+\boldsymbol{y}\|\leqslant\|\boldsymbol{x}\|+\|\boldsymbol{y}\|$（三角形不等式）。(3)的验证可根据著名的 Cauchy-Schwarz 不等式（证明见第三章）。

$$|(\boldsymbol{x}, \boldsymbol{y})|\leqslant(\boldsymbol{x}, \boldsymbol{x})^{1/2}(\boldsymbol{y}, \boldsymbol{y})^{1/2}。$$

由此又可定义非零向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的夹角 θ，由

$$\cos\theta=(\boldsymbol{x}, \boldsymbol{y})/(\boldsymbol{x}, \boldsymbol{x})^{1/2}(\boldsymbol{y}, \boldsymbol{y})^{1/2},\ 0\leqslant\theta\leqslant\pi$$

给出。

当 $(\boldsymbol{x}, \boldsymbol{y})=0$，$\theta=\frac{\pi}{2}$，称 $\boldsymbol{x}$，$\boldsymbol{y}$ **正交**，或记为 $\boldsymbol{x}\perp\boldsymbol{y}$。我们将认为零向量与任何向量正交，这样在推理时较为方便。

注意，正交概念是依赖于内积的。R^n 中除标准内积外，还可以定义其他内积（参看本章问题 9）。但在未予声明时，我们将默认正交性是在标准内积下给定的。

定义 1.7 设 S、T 是内积空间 L 的两个子集。如果

$$(\boldsymbol{x}, \boldsymbol{y})=0 \quad \forall\boldsymbol{x}\in S,\ \boldsymbol{y}\in T,$$

则称 S 与 T 正交，记为 $S\perp T$。如果

$$(\boldsymbol{x}_1, \boldsymbol{x}_2)=0 \quad \forall\boldsymbol{x}_1,\ \boldsymbol{x}_2\in S \text{ 且 } \boldsymbol{x}_1\neq\boldsymbol{x}_2,$$

则称 S 是**正交集**。若正交集 S 满足

$$\|\boldsymbol{x}\|=1 \quad \forall\boldsymbol{x}\in S,$$

称 S 是**正交规范集**。

内积空间中一个基本的事实是：可以从任一线性无关集 A 出发，去构造与 A 等价的正交规范集 S。所用的方法来源于立体几何中的正交投影的思想，称作 Schmidt **正交化方法**，见如下的

定理 1.3 设 $A=\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r\}$ 是 L 中无关集，则存在

与 A 等价的正交规范集 $S=\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_r\}$，且满足 $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k\}$ 与 $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_k\}$ 等价，$k=1, \cdots, r$。

证明 对 r 用归纳法。当 $r=1$，必有 $\boldsymbol{a}_1\neq\boldsymbol{0}$，令 $\boldsymbol{e}_1=\boldsymbol{a}_1/\|\boldsymbol{a}_1\|$，得命题成立。现设 $r=l-1$ 时命题为真，考虑 $r=l$ 时的情形：令

$$\boldsymbol{b}_l=\boldsymbol{a}_l-\sum_{i=1}^{l-1}\xi_i\boldsymbol{e}_i,\ \xi_i \text{ 待定。}$$

注意到 $(\boldsymbol{b}_l, \boldsymbol{e}_i)=0$ 即

$$\begin{aligned}0&=\left(\boldsymbol{a}_l-\sum_{i=1}^{l-1}\xi_i\boldsymbol{e}_i,\ \boldsymbol{e}_i\right)\\&=(\boldsymbol{a}_l,\ \boldsymbol{e}_i)-\left(\sum_{i=1}^{l-1}\xi_i\boldsymbol{e}_i,\ \boldsymbol{e}_i\right)=(\boldsymbol{a}_l,\ \boldsymbol{e}_i)-\xi_i,\\&\Rightarrow\xi_i=(\boldsymbol{a}_l,\ \boldsymbol{e}_i),\ i=1,\ \cdots,\ l-1。\end{aligned}$$

且由归纳假设 $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_{l-1}\}$ 可被 $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{l-1}\}$ 表出，而 $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_l\}$ 是无关的，则 $\boldsymbol{b}_l$ 不能为零。由此可令 $\boldsymbol{e}_l=\boldsymbol{b}_l/\|\boldsymbol{b}_l\|$，得 $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_l\}$ 满足定理的要求。▌

推论 维数不为零的有限维线性空间一定有正交规范基。▌

定义 1.8 设 L 是内积空间，$S\subset L$。令

$$S^\perp=\{\boldsymbol{x}:\ \boldsymbol{x}\perp S\},$$

则易证 $S^\perp$ 是 L 的子空间(思考题)，称作 S 的**正交补空间**。

习　题

1.5　证明：$S\subset(S^\perp)^\perp$。并且

$$S=(S^\perp)^\perp\Leftrightarrow S \text{ 是子空间。}$$

1.6　设 S、T 是子空间，证明 $S\subset T\Leftrightarrow S^\perp\supset T^\perp$。

1.7　如果 L 的子空间 S 与 T 的交为零向量，且 $S\perp T$，则称和 $S+T$ 为**正交直和**，记为 $S\dotplus T$。证明：当 S 是 L 的

子空间，有

$$S \dotplus S^{\perp} = L。$$

这正是称 $S^{\perp}$ 为 S 的正交补的由来。

1-3 线性映射及其矩阵表示

先引进矩阵的乘法：设 $A=[a_{ij}]$ 是 $m\times n$ 阶阵，$B=[b_{ij}]$ 是 $n\times k$ 阶阵，令

$$AB \triangleq C=[C_{ij}]$$

由 $C_{ij}=\sum_{l=1}^{n} a_{il}b_{lj}$ 给定，$i=1, \cdots, m$；$j=1, \cdots, k$。C 称作 A 与 B 的**乘积**，为 $m\times k$ 阶阵。

应当注意：要两个矩阵的乘法可行，要求（且仅要求）前一矩阵的列数等于后一矩阵的行数，因此，交换次序后乘法就可能无定义。即令 AB、BA 都可乘，也不一定有 $AB=BA$。故乘法交换律不成立，这是与数的乘法不同之处。但读者将不难验证，只要若干个矩阵的乘积有意义，乘法结合律总满足（思考题）。

一个 $n\times n$ 阶矩阵亦称 n 阶**方阵**。方阵

$$I=[\delta_{ij}]，其中\ \delta_{ij}=\begin{cases}1, & 当\ i=j,\\ 0, & 当\ i\neq j,\end{cases}$$

称为**单位阵**。如需记出它的阶数 n，就记为 I_n。I 被称为单位阵的理由是：用它去左乘任一 $n\times l$ 阶阵或去右乘任一 $k\times n$ 阶阵，并不改变该被乘的矩阵。元素全部是零的矩阵，称为**零矩阵**，仍记为 0。显然，零矩阵与任何和它可乘的矩阵的积为零矩阵。元素 a_{ii}，$i=1, \cdots, n$ 称作 A 的**主对角元**，它们所处的位置叫 A 的**主对角线**。除主对角元外全为零的方阵，称作**对角阵**，记为 $\mathrm{diag}(a_{11}, \cdots, a_{nn})$。主对角线以下（上）

全为零的矩阵称为**上(下)三角阵**。

必须指出：由 $AB=0$，不一定能推出 $A=0$ 或 $B=0$。读者可从2阶方阵中举出反例。这是矩阵乘法不同于数的乘法的又一特征(有零因子)。

如果矩阵 A 与 B 可乘，且

$$AB=I_m。$$

则称 A 是 B 的**左逆**，B 是 A 的**右逆**。这时必定有 A 为 $m\times n$ 阵而 B 为 $n\times m$ 阵，且 $n\geqslant m$。(思考题) 如果有 $n=m$，则有 $AB=BA=I$，称 B 是 A 的**逆矩阵**，记为 A^{-1}。以后将指出，A 的逆矩阵是唯一的。并且 $(A^{-1})^{-1}=A$。但是，决不是任何方阵皆有逆矩阵。我们称具有逆矩阵的方阵为**非奇异阵**，**或可逆阵**；反之，则为**奇异阵**，**不可逆阵**。

如果 n 阶实方阵 C 满足 $C^TC=CC^T=I_n$，称 C 为**正交阵**。如果 n 阶复方阵 U 满足 $U^*U=UU^*=I_n$，则称 U 为**酉阵**。

当我们把加法和乘法用于分块矩阵时，只要剖分阶数适当，所剖分的子阵可以整体地参加运算。例如：两个阶数相同的矩阵 A 与 B 都剖分为 $p\cdot q$ 块，如

$$A=\begin{bmatrix} A_{11} & \cdots & A_{1q} \\ \vdots & & \vdots \\ A_{p1} & \cdots & A_{pq} \end{bmatrix},\ B=\begin{bmatrix} B_{11} & \cdots & B_{1q} \\ \vdots & & \vdots \\ B_{p1} & \cdots & B_{pq} \end{bmatrix}。$$

并且相应的 $A_{\alpha\beta}$ 与 $B_{\alpha\beta}$ 有相同的阶数，则

$$A+B=\begin{bmatrix} A_{11}+B_{11} & \cdots & A_{1q}+B_{1q} \\ \vdots & & \vdots \\ A_{p1}+B_{p1} & \cdots & A_{pq}+B_{pq} \end{bmatrix};$$

又若 A 是 $m\times n$ 阶阵，B 是 $n\times k$ 阶阵，A 剖分为 $p\cdot q$ 块(有 p 个块行，q 个块列)，B 剖分为 $q\cdot r$ 块，且 A 剖分后的第