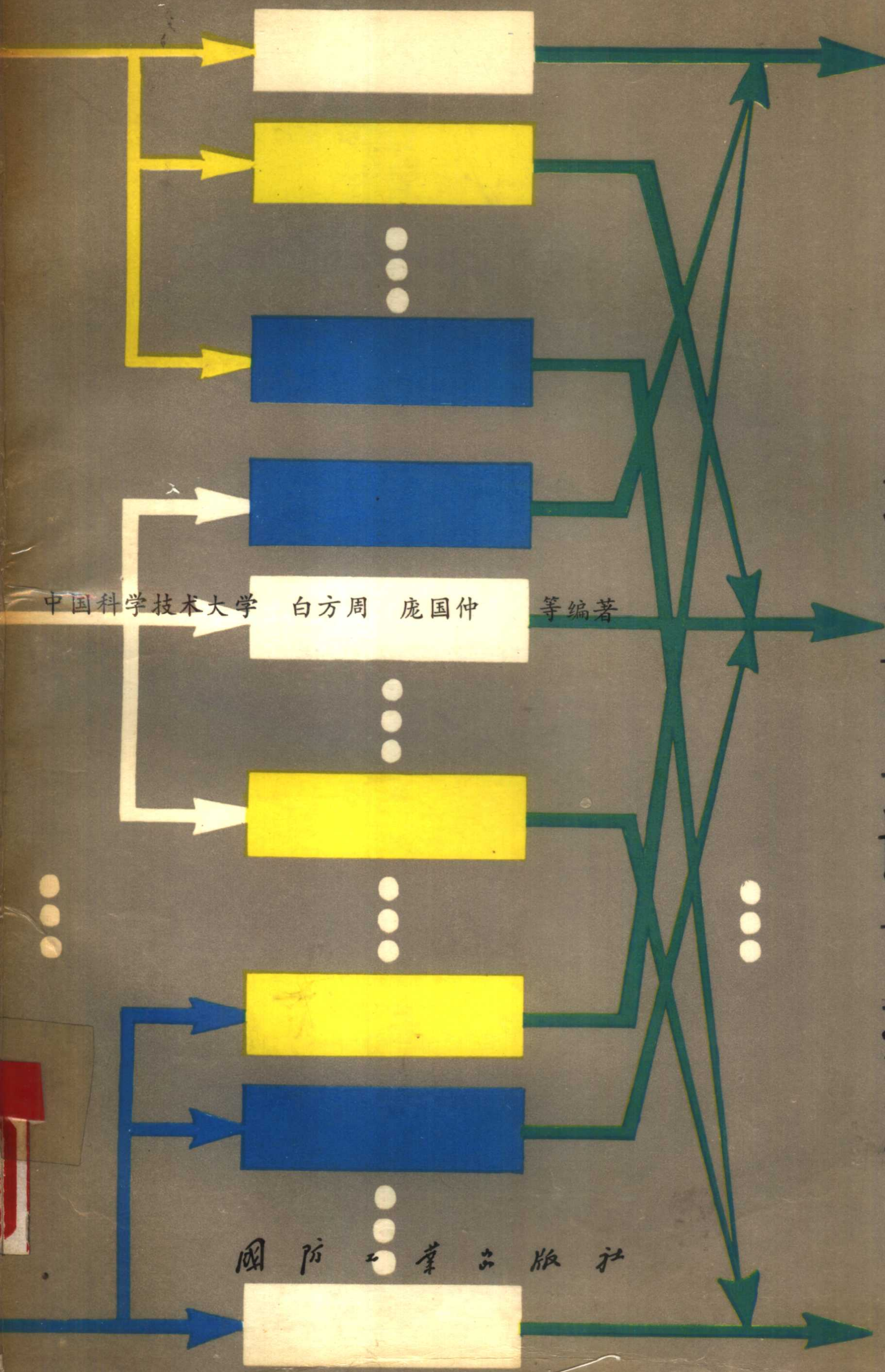
# 多变量频域理论与设计技术



中国科学技术大学　白方周　庞国仲　等编著

国防工业出版社

# 多变量频域理论与设计技术

中国科学技术大学

白方周 庞国仲 等编著

国防工业出版社

## 内容简介

七十年代初开始发展起来并日趋完善的多变量频域理论与设计技术，近几年来受到我国控制理论界和控制工程界的广泛重视。本书在介绍了多变量频域理论的有关基本概念之后，重点讨论了四种具体设计方法，即奈奎斯特阵列法、特征轨迹法、并矢展开法和序列回差法，并给出了这些方法的应用实例。鉴于采用上述设计方法完成多变量系统设计的过程离不开计算机辅助设计（CAD），所以本书以一章的篇幅专门介绍了作者和他们的课题组基于多变量频域理论开发的一个CAD软件包的概貌，并在阐述四种设计方法的后面给出了各方法的主程序清单。本书最后还讨论了关于近似模型的频域设计方法及频域鲁棒性分析问题。

本书可供从事自动控制专业的工程技术人员，特别是设计人员参考，也可作为工科院校自动控制专业的高年级学生、研究生和教师的参考书。

**多变量频域理论与设计技术**

中国科学技术大学

白方周 庞国仲 等编著

责任编辑：张均[illegible]

国防工业出版社出版

新华书店北京发行所发行　各地新华书店经售

国防工业出版社印刷厂印装

$787 \times 1092 \ 1/16$　印张$28\frac{1}{2}$　664千字

1988年2月第一版　1988年2月第一次印刷　印数：0,001— 1,500册

ISBN 7-118-00001-9/TP1　定价：6.35元

# 前　言

近二十年来，随着工业生产规模和连续生产量的不断扩大，能源和原材料价格的急剧上涨，对控制系统品质的要求不断提高，对系统可靠性和对防止环境污染的要求愈来愈严格，出现了许多复杂而困难的控制问题。其中，多变量控制就是控制理论界和控制工程界最为关注的研究的课题之一。五十年代末开始发展起来的状态空间理论，虽然在航天和航空领域获得了卓有成效的应用，然而，这种先进的时域方法并没有给从事地面上一大类复杂工业过程控制的工程师和设计师带来多大方便，追究其原因主要有：

（1）大多数工业受控对象的数学模型难以精确得到；

（2）规定这些受控对象行为的性能指标远不象宇航问题那样能以明显的形式表达出来；

（3）直接采用最优控制、最优滤波综合技术得到的控制器，其结构过于复杂，甚至不是物理可实现的，从而很难为控制工程师所接受。

于是，还在状态空间法蓬勃发展时期，不少控制理论学者就恢复了对频域法的兴趣。六十年代中期，卡尔曼（Kalman）探讨并提出了最优控制问题的频域描述，迈开了填补先进时域法和经典频域法之间缝隙的第一步。然而，最值得提出的是英国的罗森布洛克（Rosenbrock），他系统地、开创性地研究了如何将单变量系统的频率响应法推广到多变量系统设计中。他的著名论文“应用逆奈奎斯特阵列法设计多变量系统”，利用矩阵对角优势概念，把一个多变量系统的设计转化为人们熟知的单变量系统的设计问题，从而为多变量系统开辟了一条崭新的设计路线。

罗森布洛克的逆奈奎斯特阵列法的成功，带来了频域法的复兴。七十年代初，相继出现了梅奈（Mayne）的序列回差法，麦克法兰（MacFarlane）的特征轨迹法和欧文斯（Owens）的并矢展开法。这些方法的共同特点是把一个多输入-多输出、回路间紧密关联的多变量系统的设计，转化为一组单变量系统的设计问题，进而可选用某一种经典方法（如奈奎斯特和伯德的频率响应法，伊万斯的根轨迹法等）完成系统的设计。由于罗森布洛克等人发展的这些方法都是单变量频域理论在某种意义上的推广，因此保留了单变量频域设计方法的一系列优点（诸如直观、物理概念清晰、容易判别设计方向、不要求特别精确的数学模型、可以充分发挥设计者的智慧和技能，设计出既满足品质要求、又是物理可实现的结构简单的控制器，等等）。这对众多熟悉单变量频域理论的人们具有很大的吸引力。

整个七十年代，被认为是频域法大有生机的十年。在这期间，英国学者在多变量频域法的理论和实践方面作了大量的工作，获得了许多有意义和有价值的结果，形成了现代控制理论中的英国学派。国外已将上述理论用于化工、石油化工、造纸、核反应堆、飞机发动机以及飞机自动驾驶仪等设备和装置中的多变量系统的分析与设计，取得了令人信服的结果。1980年，日本统计了他们国内80家化工、石油化工等大型企业中过程控制设备（大多是多输入-多输出系统）的情况，资料表明：49％的企业采用多变量控

制方案，其中采用解耦控制的占 40%，采用最优控制的占 27%，而采用多变量频域法（逆奈奎斯特阵列法等）的占 33%。这就说明，多变量频域理论在日本工业过程控制中的应用已经相当普及了。

这里特别要指出的是，频域设计方法要依赖大量的图形信息，设计过程要以交互方式进行，而最终设计结果往往需要反复多次才能获得。这些特点表明，利用多变量频域方法完成系统的设计，其数值计算量和绘图工作量是很大的。不言而喻，这是用人工难以完成的。因此，计算机辅助设计（CAD）从一开始就是多变量频域理论研究的重要组成部分。国外几乎所有从事多变量频域理论研究并获得成就的机构和个人，都开发了不少CAD 软件包。例如：英国曼彻斯特大学理工学院的 UMIST 包；英国剑桥大学的 CAPTAIN 包；西德鲁尔大学的KEDDC 包；匈牙利布达佩斯大学的 MERCEDES 包；瑞典隆德大学的 SYNPAC 包；瑞士朱利希大学的 INTPOS 包；加拿大亚伯特大学的 GEMSCOPE 包；日本东京工业大学的DPACS-F包；日本川漪重工业公司的 CSDP-1 及CSDP-2 包；等等。在这些软件包中，有的主要是依据多变量频域理论研制的，有的则是覆盖了包括多变量频域法在内的控制理论的许多分支。

就国内而言，1980 年以来一些大学和研究机构先后开始了这一理论的研究，在全国性的一些学术会议和若干有关刊物上陆陆续续发表了不少这方面的文章、报告和连载专题讲座；嗣后不久，包括中国科技大学在内的一些学校开设了“多变量频域理论和计算机辅助设计”课程；接着，基于这种理论的程序包出现了。目前不少单位正将这一理论推进到应用阶段，并已开始见到实效。

以上情况表明，国内对多变量频域理论及其应用感兴趣的人越来越多了。为了进一步推广这一理论，适应日益增多的从事多变量频域理论研究的科技人员和从事近代工业过程控制的工程师们的需要，我们编著了这本书。本书是这样构成的：第一章提供阅读本书所必要的数学基础；第二章扼要回顾了单变量控制理论，重点是经典频域法的基本内容；第三章阐述了多变量系统的描述形式和若干基本概念，其中包括罗森布洛克（Rosenbrock）提出的系统矩阵的描述形式和解耦零点等概念；第四章首先通过对几个典型多变量工业装置的讨论，给出了多变量系统的一般结构形式和基本数学关系式，而重点是分析多变量系统的基本性能指标（稳定性、关联性、整体性和稳态精确度等）；第五章至第八章分别介绍了几种主要的多变量频域设计方法，即奈奎斯特阵列法、特征轨迹法、并矢展开法和序列回差法，这四章的末尾都附有相应于各种设计方法的主程序清单，供有兴趣的读者和具体从事多变量系统设计者参考；第九章给出了几个应用上述设计方法的实例，其中有国外文献上刊载的，有作者亲自在国内大型工业设备上实践过的；第十章介绍了我们基于多变量频域理论开发的一个 CAD 软件包——MFCSDP 的概貌，其内容包括软件包的设计思想、程序功能、软件结构和特点，以及操作和维护的说明等，最后用一节讨论了MFCSDP 的系统仿真问题。第十一章介绍近似模型设计技术，这里面给出了国内外学者和我们自己新近的研究成果，它是控制理论和控制工程界十分关注并有强烈工程背景的课题之一。本章的重点放在近似模型的频域设计方法和频域鲁棒性分析上。

本书第一、三、五章，以及第九章的前两节由白方周同志编写；第二、六、七章，以及第九章的后两节由庞国仲同志编写；第四章由庞国仲和李嗣福同志编写；第八章由

赵守忠同志编写；第十章由濮洪钧同志编写；第十一章由鲍远律同志编写。全书由白方周、庞国仲同志主编。

中国科技大学十系的李高文、杨泽民、薛福珍同志以及研究生薛建平、罗宁苏、杨生虎和李捷同志在程序编写和系统设计方面作了很多工作，编者在此向他们表示衷心感谢！

由于编者业务水平所限，书中一定存在不少缺点和错误，敬请读者批评指正。

# 目　录

# 第一章　必要的数学基础

本章扼要地介绍一些与控制系统理论有关的线性代数基本知识，其中包括：向量空间的基本概念；常数矩阵的定义和某些结果；行列式；特征值和特征向量；并矢展开式；格希高林（Gershgorin）定理和奥斯佐夫斯基（Ostrowski）定理；矩阵的对角优势概念；多项式和多项式矩阵等有关内容。在叙述中，除个别重要定理外我们均略去证明，读者如有必要，可查阅普通的线性代数和矩阵理论书。

## §1.1　向量空间

### 1．集合

我们把某些具有共同内在属性的事物称作集合，集合中的每一件事物称作元素（或简称为元）。通常，以大写字母$A$，$B$，$C$，…，$X$，$Y$表示集合，而用小写字母$\boldsymbol{a}$，$\boldsymbol{b}$，$\boldsymbol{c}$，…，$\boldsymbol{x}$，$\boldsymbol{y}$表示相应集合中的元素。设$X$是一个集合，$\boldsymbol{x}$是$X$中的一个元素，则

$$\boldsymbol{x}\in X$$

表示元素$\boldsymbol{x}$属于集合$X$；若$\boldsymbol{x}$不属于$X$，可表示为：

$$\boldsymbol{x}\notin X$$

若$A$为一给定的集合，集合$B$是$A$的一部分，则称$B$为$A$的子集合（或子集），记为

$$B\subset A 或 A\supset B，$$

意思为“$B$含于$A$”，或“$A$包含$B$”。如果两个集合$A$和$B$同时满足$A\subset B$及$B\subset A$，则$A$和$B$相同，记为

$$A=B$$

如果集合$A$中包含有无穷多个元，则称$A$为无穷集；而若集合$A$中只含有有限多个元时，则称$A$为有限集或有穷集。一般由$n$个元组成的有限集常记为

$$\{\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \cdots,\ \boldsymbol{a}_n\}$$

不包含任何事物的集合叫作空集，空集用$\varnothing$表示。

### 2．向量空间

设在集合$X$中定义了加法和数乘运算，那么，对于集合$X$中的任意元$\boldsymbol{x}$，$\boldsymbol{y}$，$\boldsymbol{z}$和实数域$R$中的纯量$\alpha$，$\beta$，如果满足

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

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

（3）对于$X$中的$\mathbf{0}$元，满足

$$\boldsymbol{x}+\mathbf{0}=\boldsymbol{x}\qquad 对所有\boldsymbol{x}\in X$$

（4）对任意的$\boldsymbol{x}\in X$，都恒有相应的负元素$-\boldsymbol{x}\in X$，满足

$$\boldsymbol{x}+(-\boldsymbol{x})=\mathbf{0}$$

（5）$\alpha(\beta\boldsymbol{x})=(\alpha\beta)\boldsymbol{x}$

(6) $\alpha(\boldsymbol{x}+\boldsymbol{y})=\alpha\boldsymbol{x}+\alpha\boldsymbol{y}$

(7) $(\alpha+\beta)\boldsymbol{x}=\alpha\boldsymbol{x}+\beta\boldsymbol{y}$

(8) 设1是$R$中的单位元，有

$$1\cdot\boldsymbol{x}=\boldsymbol{x}\quad 对所有的\boldsymbol{x}\in X$$

则称$X$为在域$R$上的向量空间或线性空间，记作$R^n$。$X$中的元称为这个线性空间的向量或点。$n$维实数向量空间$R^n$就是$n$元实数组的集合。

设$R^n$中的两个向量$\boldsymbol{x}$和$\boldsymbol{y}$

$$\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\quad \boldsymbol{y}=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}$$

加法规则定义为

$$\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}+\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}=\begin{pmatrix}x_1+y_1\\x_2+y_2\\\vdots\\x_n+y_n\end{pmatrix}\in R^n$$

向量和实数$\alpha$的乘法（亦称数乘）定义为

$$\alpha\begin{pmatrix}x_1\\x\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}\alpha x_1\\\alpha x_2\\\vdots\\\alpha x_n\end{pmatrix}\in R^n$$

$n$维复数向量空间$C^n$可以用类似的方法加以定义。

**3. 向量组的线性相关性**

引入向量组的线性相关性、线性无关性的概念，是为了研究一组向量间的关系。实际上，这一概念是向量组的“平行”、“共线”、“共面”等概念的概括和扩充。设$X$是$R$上的线性空间，$\boldsymbol{x}_1$，$\boldsymbol{x}_2\cdots$，$\boldsymbol{x}_n$是$X$中的一组向量，如果在$R$中存在不全为零的$n$个数$k_1$、$k_2$，…，$k_n$，满足

$$k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2+\cdots+k_n\boldsymbol{x}_n=\boldsymbol{0}\in X \tag{1.1.1}$$

则称向量$\boldsymbol{x}_1$，$\boldsymbol{x}_2$，…，$\boldsymbol{x}_n$为线性相关；如果找不到一组不全为零的$n$个数$k_1$，$k_2$，…，$k_n$，或者说，由等式（1.1.1）只能推出$k_1=k_2=\cdots=k_n=0$，则称此向量组为线性无关。由式（1.1.1）不难导出向量组线性相关的充分必要条件为：其中必有一向量可以表示成其余向量的线性组合。例如在式（1.1.1）中，若$k_1\neq 0$，则有

$$\boldsymbol{x}_1=-\frac{k_2}{k_1}\boldsymbol{x}_2-\frac{k_3}{k_1}\boldsymbol{x}_3\cdots-\frac{k_n}{k_1}\boldsymbol{x}_n \tag{1.1.2}$$

如果在线性空间$X$中有$n$个线性无关的向量，但是没有更多数目的线性无关向量，那么就称$X$为$n$维的，记作

$$dim X=n$$

换言之，$n$维空间中线性无关向量的最大个数为$n$。

**4. 线性空间的基和坐标**

设$\boldsymbol{x}_1$，$\boldsymbol{x}_2$，…，$\boldsymbol{x}_n$为$n$维线性空间$X$中的一个有序向量组，$k_1$，$k_2$，…，$k_n$为数域$R$中的一个有序数组，如果

（1）$\boldsymbol{x}_1$，$\boldsymbol{x}_2$，…，$\boldsymbol{x}_n$是线性无关的；

（2）$X$中的任意向量$\boldsymbol{x}$都可以表示成$\boldsymbol{x}_1$，$\boldsymbol{x}_2$，…，$\boldsymbol{x}_n$的线性组合

$$\boldsymbol{x}=k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2+\cdots+k_n\boldsymbol{x}_n \tag{1.1.3}$$

则把向量组$\boldsymbol{x}_1$，$\boldsymbol{x}_2$，…，$\boldsymbol{x}_n$称为向量空间$X$中的一组基，每一$\boldsymbol{x}_i$称为基向量，式（1.1.3）称为$\boldsymbol{x}$在基$\boldsymbol{x}_1$，$\boldsymbol{x}_2$，…，$\boldsymbol{x}_n$上的分解式。式中（$k_1$，$k_2$，…，$k_n$）被向量$\boldsymbol{x}$和$\boldsymbol{x}_1$，$\boldsymbol{x}_2$，…，$\boldsymbol{x}_n$唯一确定，这个有序数组就称为$\boldsymbol{x}$在基$\boldsymbol{x}_1$，$\boldsymbol{x}_2$，…，$\boldsymbol{x}_n$上的坐标。

在基$\boldsymbol{x}_1$，$\boldsymbol{x}_2$，…，$\boldsymbol{x}_n$下，向量$\boldsymbol{x}_i$本身当然也有坐标，用$\boldsymbol{e}_i$表示$\boldsymbol{x}_i$在这一组基上的坐标，显然有

$$\boldsymbol{e}_i=\begin{array}{c}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{array}\leftarrow\text{第 } i \text{ 个元素},$$

称$\boldsymbol{e}_i$（$i=1$，$2$，…，$n$）为单位向量，而称向量组$\boldsymbol{e}_1$，$\boldsymbol{e}_2$，…，$\boldsymbol{e}_n$为自然基或标准基。

## §1.2 矩　　阵

**1. 矩阵定义**

将$m\times n$个元素（可以是实数，复数，函数…）作如下排列

$$A=\{a_{ij}\}_{m\times n}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix} \tag{1.2.1}$$

定义$A$为矩阵。式中$a_{ij}$表示矩阵中第$i$行、$j$列上的元素，$m$代表行数，$n$代表列数，通常$m\neq n$；当$m=n$时，矩阵$A$为一方阵，它的阶数为$n$。

当$n=1$，即只有一列时：

$$\begin{pmatrix} a_1\\ a_2\\ \vdots\\ a_m\end{pmatrix}$$

为$m\times 1$阵，又称为$m$维列向量。

当$m=1$时，即只有一行时：

$$(a_1,\ a_2,\ \cdots,\ a_n)$$

为$1\times n$阵，又称为$n$维行向量。

矩阵的基本运算，如矩阵加法，矩阵和标量的数乘以及矩阵乘法等，读者都是熟知的，本书在此就不赘述了。

**2．转置和共轭转置矩阵**

$m\times n$ 矩阵 $A$ 的转置记为 $A^T$，它是依次把 $A$ 的行改写成列而得到的：

$$A^T=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}^T=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix} \tag{1.2.2}$$

显然，如果 $A$ 是 $m\times n$ 阵，那么 $A^T$ 是 $n\times m$ 阵。转置矩阵的重要性质有：

（1）$(A^T)^T=A$； (1.2.3)

（2）对于方阵 $A$，如果 $A^T=A$，即 $a_{ij}=a_{ji}$，则称 $A$ 为对称矩阵；

（3）对于矩阵 $A$，$B$，有

$$(AB)^T=B^TA^T \tag{1.2.4}$$

对于复数域 $C$ 上的矩阵 $A$，如果将 $A$ 的各元素 $a_{ij}$ 用其共轭复数 $\bar{a}_{ij}$ 置换，则得到的矩阵 $(\bar{a}_{ij})$ 记为 $\bar{A}$，称为 $A$ 的共轭阵。$A$ 的共轭转置记为 $A^+$，且有下列性质：

（1）当 $A$ 为实阵时，$A^+=A^T$；

（2）$(AB)^+=B^+A^+$；

（3）对于方阵，当 $A=A^+$ 时，称 $A$ 为厄米特（Hermitian）矩阵。任意厄米特矩阵 $A$ 可以表示成

$$A=A_1+jA_2 \tag{1.2.5}$$

其中 $A_1$，$A_2$ 都是厄米特矩阵。

**3．矩阵的秩**

设 $A$ 为一 $m\times n$ 矩阵：

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{1.2.6}$$

这时，$A$ 共有 $n$ 个 $m$ 维列向量或者 $m$ 个 $n$ 维行向量。如果 $A$ 的 $n$ 个列向量中有 $r$ 个线性无关（$r<n$），而任意 $r+1$ 个列向量均线性相关，则定义整数 $r$ 为矩阵 $A$ 的秩，记作

$$\operatorname{rank}A=r$$

显然，$r\leqslant n$。此时，矩阵 $A$ 的 $m$ 个行向量也只有 $r$ 个是线性无关的，因此又有 $r\leqslant m$，于是

$$\operatorname{rank}A=r\leqslant\min\{m,\ n\} \tag{1.2.7}$$

如果矩阵 $A$ 的秩等于该阵的列数，则称 $A$ 为列满秩；如果 $A$ 的秩等于该阵的行数，则称 $A$ 为行满秩，一般都不加区分地统称为满秩。可以证明

$$\operatorname{rank}A=\operatorname{rank}A^T \tag{1.2.8}$$

$$\operatorname{rank}(A+B)\leqslant\operatorname{rank}A+\operatorname{rank}B \tag{1.2.9}$$

$$\operatorname{rank}(AB)\leqslant\min(\operatorname{rank}A,\ \operatorname{rank}B) \tag{1.2.10}$$

**4．矩阵的迹**

$n$ 阶方阵 $A$ 的迹定义为

$$\operatorname{tr} A = a_{11} + a_{22} + \cdots + a_{nn} \tag{1.2.11}$$

容易证明：如果 $A = \{a_{ij}\}_{n\times m}$，$B = \{b_{ij}\}_{m\times n}$，则有

$$\operatorname{tr}(AB) = \operatorname{tr}(BA) \tag{1.2.12}$$

**5. 逆矩阵**

设 $A$ 为 $n$ 阶方阵，如果存在另一个 $n$ 阶方阵 $B$，满足

$$AB = BA = I_n$$

其中 $I_n$ 为 $n \times n$ 单位矩阵，则称 $B$ 为 $A$ 的逆矩阵，记为

$$B = A^{-1} = \hat{A}$$

由于

$$\operatorname{rank} I_n = n = \operatorname{rank} AA^{-1} \leqslant \min(\operatorname{rank} A,\ \operatorname{rank} A^{-1}) \leqslant n \tag{1.2.13}$$

因此可以看出，仅当 $\operatorname{rank} A = n$ 时，$A^{-1}$ 才可能存在，这个条件也是充分条件。

**6. 矩阵的分块形式**

一个 $m \times n$ 矩阵中含有 $m \times n$ 个元素，为方便起见，可以把它表示成分块形式：

$$A = \begin{pmatrix} A_{11} & \cdots\cdots & A_{1l} \\ \vdots & & \vdots \\ \underbrace{A_{q1}}_{n_1} & \cdots\cdots & \underbrace{A_{ql}}_{n_l} \end{pmatrix} \begin{matrix} \}m_1 \\ \\ \}m_q \end{matrix} \tag{1.2.14}$$

这时，$A$ 被分成 $q \times l$ 块，每个 $A_{ij}$ 都是一个矩阵（子阵），其维数为 $m_i \times n_j$，它当然可以是 $1 \times 1$ 维的，也可以是列向量或行向量。

值得指出的是，一个矩阵的分块并不是唯一的；而且，分块前后的运算规则相同。

分块矩阵的相乘可以表示为

$$C = AB = \begin{pmatrix} A_{11} & \cdots & A_{1r} \\ \vdots & & \vdots \\ A_{q1} & \cdots & A_{qr} \end{pmatrix} \begin{pmatrix} B_{11} & \cdots & B_{1l} \\ \vdots & & \vdots \\ B_{r1} & \cdots & B_{rl} \end{pmatrix} = \begin{pmatrix} C_{11} & \cdots & C_{1l} \\ \vdots & & \vdots \\ C_{q1} & \cdots & C_{ql} \end{pmatrix} \tag{1.2.15}$$

其中

$$C_{ij} = \sum_{k=1}^{r} A_{ik} B_{kj} \tag{1.2.16}$$

在式（1.2.14）中，如果 $A_{ij} = 0$（$i \neq j$，$q = l$），则称 $A$ 为分块对角阵，即

$$A = \begin{pmatrix} A_{11} & & 0 \\ & A_{22} & \\ & & \ddots \\ 0 & & A_{ll} \end{pmatrix} \tag{1.2.17}$$

式中 $A_{11}$，$A_{22}$，…、$A_{ll}$ 可以不是方阵。

如果分块后的 $A$ 为

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1l} \\ & \ddots & \\ 0 & & A_{ll} \end{pmatrix} \text{或} \quad A = \begin{pmatrix} A_{11} & & 0 \\ \vdots & \ddots & \\ A_{l1} & \cdots & A_{ll} \end{pmatrix} \tag{1.2.18}$$

则称 $A$ 为上三角分块阵或下三角分块阵。当 $l = 2$，并且 $A_{11}$，$A_{22}$ 是可逆的，则有以下

两个计算三角分块阵的逆阵公式：

（1）$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{pmatrix} \tag{1.2.19}$$

（2）$$\begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A_{11}^{-1} & 0 \\ -A_{22}^{-1}A_{21}A_{11}^{-1} & A_{22}^{-1} \end{pmatrix} \tag{1.2.20}$$

## §1.3 矩阵的行列式

当矩阵$A$为一方阵时，就存在有行列式，通常记为

$$\det A = |A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \tag{1.3.1}$$

行列式是矩阵的一种数字表征。由行列式定义可推得如下基本性质：

（1）任意交换两相邻的行或列，行列式只改变符号。

（2）如果任意两行（列）的元素对应相等，则行列式的值为零。

（3）如果任意一行（列）的元素加上另一行（列）元素的常数倍，则行列式的值保持不变。

（4）$\det A^T = \det A$ (1.3.2)

（5）$\det(kA) = k^n \det A$ (1.3.3)

（6）$\det AB = \det BA = \det A \det B$ (1.3.4)

（7）$\det[\operatorname{diag}\{\lambda_i\}_{1\leq i\leq n}] = \lambda_1\lambda_2\cdots\lambda_n$ (1.3.5)

（8）如果方阵$A$是一个对角分块阵或三角分块阵，则

$$\det A = \det A_{11} \cdot \det A_{22} \cdots \det A_{ll} \tag{1.3.6}$$

行列式可按行或列展开。当按第$i$行展开时

$$\det A = \sum_{j=1}^{n} a_{ij}A_{ij} = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} \tag{1.3.7}$$

式中，$A_{ij}$是划去$a_{ij}$所在的第$i$行和第$j$列后得到的（$n-1$）阶矩阵行列式再乘以$(-1)^{i+j}$，叫作$a_{ij}$的代数余子式。如果用$a_{ij}$乘以其它行的代数余子式$A_{kj}$($k \neq i$)，则有$\sum_{j=1}^{n} a_{ij}A_{kj} = 0$。当按第$j$列展开时

$$\det A = \sum_{i=1}^{n} a_{ij}A_{ij} = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} \tag{1.3.8}$$

同样，若用$a_{ij}$乘以其它列的代数余子式$A_{ik}$（$k \neq j$），则有$\sum_{i=1}^{n} a_{ij}A_{ik} = 0$。

下面介绍几个常用公式：

（1）矩阵求逆公式　若$\det A \neq 0$，则称$A$为可逆矩阵（或满秩矩阵，非奇异矩阵）；反之，若$\det A = 0$，则称$A$为不可逆矩阵（或降秩矩阵，奇异矩阵）。$A$的逆阵

记为$A^{-1}$，可用下式来求：

$$A^{-1}=\frac{\mathrm{adj}\,A}{\det A} \tag{1.3.9}$$

式中

$$\mathrm{adj}\,A=\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} \tag{1.3.10}$$

称为$A$的伴随矩阵。

（2）修耳（Schur）公式　设$A$可写成下列分块形式

$$A=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

其中$A_{11}$为非奇异方阵。考虑等式

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

并取其行列式，可得如下期望的结果：

$$|A|=|A_{11}|\cdot|A_{22}-A_{21}A_{11}^{-1}A_{12}| \tag{1.3.11}$$

这是后面常用到的一个公式。

（3）设$A$和$B$分别为$m\times n$和$n\times m$矩阵，则有

$$|I_m+AB|=|I_n+BA| \tag{1.3.12}$$

此式亦非常有用（证明从略），特别当$\alpha$，$\beta$为$n$维列向量时，可得

$$|I_n+\alpha\beta^T|=1+\beta^T\alpha \tag{1.3.13}$$

如果$\beta^T\alpha\neq-1$，则矩阵$A=I_n+\alpha\beta^T$为非奇异的，并且容易证明

$$A^{-1}=I_n-\frac{1}{1+\beta^T\alpha}\alpha\beta^T \tag{1.3.14}$$

## §1.4　特征值和特征向量

### 1. 特征值、特征向量的定义及性质

方阵最重要的概念之一就是它们的特征值和特征向量。由于实系数多项式一般有复数根，因此，下面涉及的向量及空间都理解为复数域$C$上的向量及空间。

设$A$为一个$n$阶方阵，记

$$\Phi(\lambda)=\det(\lambda I-A)=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0 \tag{1.4.1}$$

$\Phi(\lambda)$是一个实系数多项式，称它为$A$的特征多项式，而把$\Phi(\lambda)=0$称为$A$的特征方程，它的根$\lambda_1$，$\lambda_2$，…，$\lambda_n$定义为$A$的特征值。特征值具有如下性质：

（1）$\prod\limits_{i=1}^{n}\lambda_i=\det A$及　(1.4.2)

$\sum\limits_{i=1}^{n}\lambda_i=\mathrm{tr}\,A$；　(1.4.3)

（2）$A'$和$A$具有相同的特征值；

（3）如果$A$是可逆矩阵，则$A^{-1}$的特征值为$1/\lambda_1$，$1/\lambda_2$，…，$1/\lambda_n$；

（4）$A-\alpha I$的特征值为$\lambda_1-\alpha$，$\lambda_2-\alpha$，…，$\lambda_n-\alpha$；

（5）若$A^k=\underbrace{A\cdot A\cdots A}_{k}$，则$A^k$的特征值为$\lambda_1^k$，$\lambda_2^k$，…，$\lambda_n^k$。

设$\lambda_i$是$A$的特征值，考察方程

$$A\boldsymbol{x}=\lambda_i\boldsymbol{x} \tag{1.4.4}$$

或写作

$$(\lambda_i I-A)\boldsymbol{x}=0 \tag{1.4.5}$$

由于$\det(\lambda_i I-A)=\Phi(\lambda_i)=0$，表明齐次方程（1.4.5）的系数矩阵不满秩，亦即表明该方程具有非零解$\xi$。换句话说，即存在非零向量$\xi$，满足

$$A\xi=\lambda_i\xi \tag{1.4.6}$$

称$\xi$为对应于$\lambda_i$的特征向量。从几何意义上说，矩阵$A$的特征向量是这样一个向量：将$A$前乘$\xi$，得到和$\xi$共线的一个向量$A\xi$，它或者与$\xi$同向，或者与$\xi$反向。

**2．相似矩阵和矩阵的标准形**

（1）设$A$，$P$和$P^{-1}$都是$n\times n$矩阵，若满足

$$\widetilde{A}=P^{-1}AP \tag{1.4.7}$$

则称$A$和$\widetilde{A}$为相似矩阵，而获得矩阵$\widetilde{A}$的运算过程，称为相似变换；把可逆矩阵$P$称为变换矩阵。可以证明，相似矩阵具有相同的特征值。

（2）矩阵的标准形　①若$A$有特征值$\lambda_1$，$\lambda_2$，…，$\lambda_n$和一组线性无关的特征向量$\boldsymbol{w}_1$，$\boldsymbol{w}_2$，…，$\boldsymbol{w}_n$，则特征向量矩阵

$$W=(\boldsymbol{w}_1,\ \boldsymbol{w}_2,\ \cdots,\ \boldsymbol{w}_n) \tag{1.4.8}$$

是非奇异的，用它可将$A$变换成对角标准形

$$\widetilde{A}=W^{-1}AW=\begin{pmatrix}\lambda_1 & & 0\\ & \lambda_2 & \\ & & \ddots \\ 0 & & \lambda_n\end{pmatrix} \tag{1.4.9}$$

②若$A$有重特征值$\lambda_i$（$i=1,\ 2,\ \cdots,\ p$），每一特征值各重复$\mu_i$次，且$\sum\limits_{i=1}^{p}\mu_i=n$，则必存在一个非奇异矩阵$W$，用它可将$A$变换成约当标准形

$$\widetilde{A}=W^{-1}AW=\begin{pmatrix}J_1 & 0 & \cdots & 0\\ 0 & J_2 & \cdots & 0\\ & \cdots\cdots & & \\ 0 & 0 & \cdots & J_p\end{pmatrix} \tag{1.4.10}$$

其中对角块为$\mu_i\times\mu_i$阵，其形式为

$$J_i=\begin{pmatrix}\lambda_i & 1 & 0 & \cdots & 0\\ 0 & \lambda_2 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & & \\ 0 & 0 & 0 & \cdots & 1\\ 0 & & \cdots\cdots & & \lambda_i\end{pmatrix} \tag{1.4.11}$$