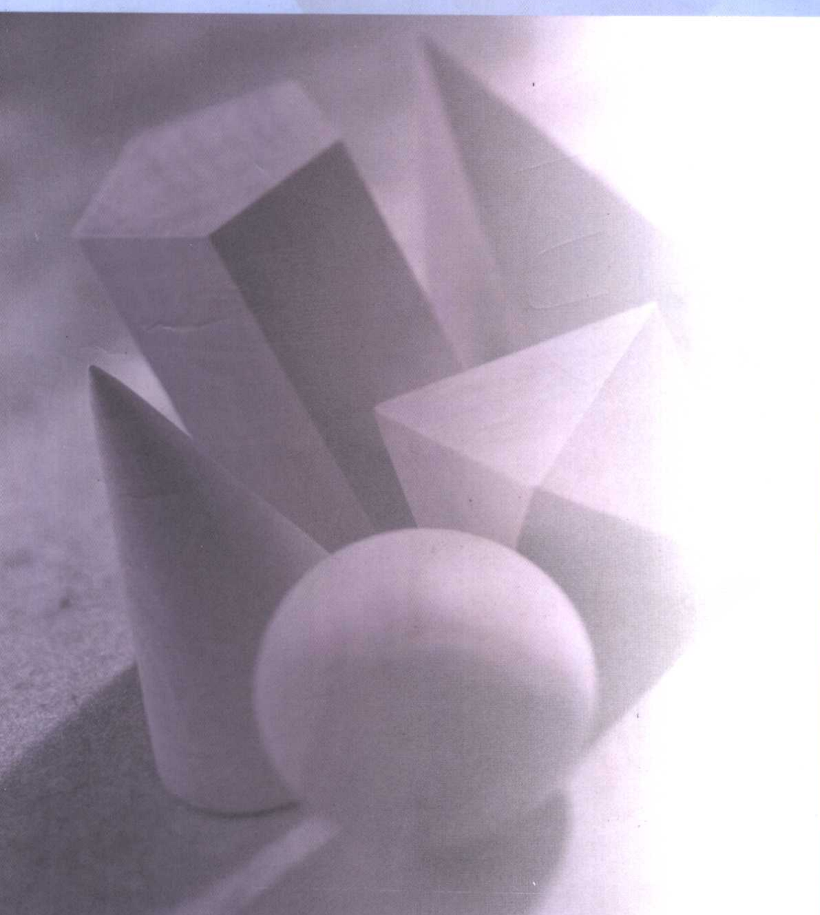
高等学校研究生教学用书

应用数值分析

（第二版）



文世鹏　编著

石油工业出版社

高等学校研究生教学用书

应用数值分析

（第二版）

文世鹏　编著

杨小远　张　明　陈安乐　编

石油工业出版社

内 容 提 要

本书系统论述了求解数学问题的数值算法的理论原理和方法。全书共分十二章，包括数值代数和数值逼近的主要内容、非线性问题的数值方法、常微分方程数值解法和 MATLAB 应用基础五个部分。书中的算法还包括有近代较新的方法，并与科学计算软件 MATLAB 的实现相结合。

作为工科研究生数值分析课的教学用书，书中内容与本科生的计算方法课程内容相衔接，力求理论与实际应用并重，循序渐进且自成体系。本书既可作为高等学校数值分析课的教材，也可供科技工程人员参考和自学。

图书在版编目(CIP)数据

应用数值分析/文世鹏编著.—2 版

北京:石油工业出版社，2001.9

高等学校研究生教学用书

ISBN 7-5021-3475-1

Ⅰ.应…

Ⅱ.文…

Ⅲ.计算方法-研究生-教材

Ⅳ.0241

中国版本图书馆 CIP 数据核字(2001)第 051273 号

石油工业出版社出版

(100011 北京安定门外安华里二区一号楼)

石油工业出版社印刷厂 排版

北京精美实华制作中心 排版

石油工业出版社印刷厂印刷

新华书店北京发行所发行

*

787×1092 毫米 16 开本 27.75 印张 693 千字 印 1—2500

2001 年 9 月北京第 2 版 2001 年 9 月北京第 1 次印刷

ISBN 7-5021-3475-1/TE·2577(课)

定价：29.80 元

再版前言

本书自出版以来，作为研究生数值分析课教材在教学中使用过多次，受到读者欢迎，并获得2000年度中国石油天然气集团公司颁发的优秀教学成果奖。根据教学需要和读者要求，决定对原书进行必要的修订补充后重新出版。

这次再版除了对原书的内容和例题有一些订正外，最大的变动是新增加了第十二章“MATLAB与数值计算”以及“MATLAB应用基础”。MATLAB作为近代优秀的功能强大而又使用方便的科学计算应用软件，已在国内外得到了广泛的应用。这两年，我们在教学中已经把数值分析课与MATLAB相结合，用MATLAB作为学习数值分析课的基本计算机语言和计算工具，取得了很好的效果。这次再版增补的内容也是我们教学实践积累的结果。

在此，要感谢我们的学生对本书再版的关心和支持鼓励，提出了很多好的建议；感谢负责再版的工作人员的努力，使得新书能顺利及时出版。

文世鹏　张　明

2001.7.30

前　　言

随着计算机科学的发展，数值计算的方法已成为自然科学、工程技术领域中解决实际问题最有效的手段。科学计算能力是现代高层次人材综合素质中所不可缺少的。而高等学校中以现代数值方法的理论原理、计算及计算机实现为主要内容的数值分析课，在培养研究生的科学计算能力方面有着重要的、不可替代的作用。因此，大多数工程技术和应用科学方面的硕士研究生都选择数值分析课作为必修的学位课。这本《应用数值分析》就是作者在为工科硕士研究生多年讲授数值分析课的基础上形成的，并根据积累的教学经验和信息，对内容进行了认真的选择增补和重新安排。

研究生数值分析课是在大学本科计算方法课的基础上的深入和提高，主要包括计算数学中数值代数和数值逼近及相关内容。在计算数学中这两部分的内容很多，而作为一本教材如何选择安排内容呢？根据多年的教学实践及国内外同类学校相关信息，我们有以下几点认识：

1. 数值分析课不应该只是罗列现有的基本算法及重复已有的程序编码流程。应该把主要精力放在提高学生分析算法，能较快接受和应用现代高效算法及构造设计专用算法的能力上。这就要求学生具有计算数学中必要的理论基础，如矩阵论基础、泛函分析的基本思想方法以及用抽象数学空间、向量和矩阵认识和处理问题的能力。

2. 工科硕士研究生学习数值分析课的主要目的，不是为了将来从事计算数学中数值代数和数值逼近本身的研究工作。因此，在数值分析课中虽然应重视加强理论基础，但不能、也没有必要过于理论化和专门化，其目的应该是让学生较好地了解一个好的数值方法是在什么理论原理的基础上产生的，以及怎样认识、比较和选择不同算法。

3. 由于计算机科学的迅速发展，不断出现一些新的数值方法。因此，在数值分析课中不能只停留在学习成熟的传统的基本数值方法上，还需要接触和了解一些近代出现的新的数值方法。

4. 要处理好与大学本科计算方法内容的衔接与提高深入。

以上几点虽然认识到了，但不一定能在实际中做得很好，我们应尽力沿着这个方向努力。全书共分十一章，内容大致分为四个部分。

第一部分：包括第一章，第二章。第一章是对大学本科线性代数重要内容的总结复习和补充深化，这是学习本书必备的基础，熟悉这部分的读者可以跳过本章。第二章是矩阵论中矩阵分解的内容，包括重要的三角分解、正交分解和奇异值分解。本书中这部分内容与矩阵论课中不同的是，不仅讨论矩阵分解的理论原理，还讨论实现这些矩阵分解，并适合在计算机上计算的具体算法。安排第一部分这两章内容，使得本书可以基本上自成体系。

第二部分：包括第三，四，五，六，十章。这些章包括了数值代数中的三个主要内容：线性代数方程组的解法，最小二乘问题和代数特征值问题。并专门安排了第六章非线性方程组的数值方法。非线性问题是当前各个领域中非常重要、非常活跃的问题。在第六章中重点讨论了基本而有效的牛顿型算法和无约束优化算法。

第三部分：包括第七，八，九章。这一部分是数值逼近的主要内容，包括函数插值、函

数最佳逼近和数值积分。在函数最佳逼近中，重点介绍了最佳平方逼近。另外，还介绍了二维区域的函数插值和曲面拟合问题。

第四部分是第十一章，常微分方程初边值问题的解法。除了初值问题外，对常微分方程边值问题也作了较多介绍，其中打靶法也包括了作者自己做过的研究课题的结果。

本书的主要对象是非计算数学专业的工科硕士研究生，具有相当于大学本科数学基础知识的工程技术人员也可以本书作为学习进修的用书，对部分工程学科的博士生也是有用的参考书。

本书的内容比实际教学的内容要多，在安排教学计划时，可以根据不同的对象和不同的要求对内容进行选择和重新组合。根据国家教委曾提出过的硕士研究生数值分析课的基本要求和作者教学实践，除去打“*”号的章节，其它的章节作为本书的主要内容，可以在60学时的课堂教学中完成。为了熟悉算法和熟悉应用数值计算软件，配合课堂理论教学可安排课后有30～40小时计算机的实习时间。

本书由文世鹏主编，并编写了第一～五章和第十章内容，承担审核修订全书。杨小远、张明、陈安乐共同参加编写工作，提供了六，七，八，九，十一章内容。在成书过程中，梁景伟、潘建华也提供过部分章节初稿，卢名高、蔡璋琏提出过许多好的建设性意见，在此表示感谢。

书中错误和不当之处，恳请读者批评指正。

编　者

1999年5月

目　录

第一章　数值分析基础

数值分析这门学科是研究求解各种数学问题的数值算法的理论和方法。学习数值分析要求有较好的线性代数知识和矩阵论初步的基础知识。许多数值方法的理论基础就是矩阵论。本章内容包括对大学阶段矩阵代数重要内容的复习和补充，以及为学习本书后面各章提供必备的矩阵论等基础知识。这些内容着重概念、结论和归纳小结，除必要外一般不作详尽的证明推导。

第一节　矩阵代数基础

一、集合的基本概念和逻辑符号

集合：集合是近代数学研究的最基本的对象，它是由一些元素组成。通常习惯上用大写字母表示集合，用小写字母表示集合中的元素，如

$a \in A$,表示 a 是集合 A 的元素(a 属于集合 A);

$b \notin A$,表示 b 不是集合 A 的元素(b 不属于集合 A)。

一个集合必须明确规定它所包含的元素,一般用列举法或充要条件描述法来表示集合。当集合中的元素是很少个数的有限集或者是可列的无限集时,可以列举其全部元素来表示集合。如集合 A 有四个元素 a,b,c,d,则可表示为 $A=\{a,b,c,d\}$。当集合中元素很多或者是无限集时,通常用刻划元素属于这个集合的充分必要条件来表示集合。其一般形式为 $X=\{x \mid p(x)\}$,它表示集合 X 是满足条件 $p(x)$ 的元素 x 的全体所组成的集合。没有任何元素的集合称为空集,用 Φ 表示。

集合可以由任意确定的事和物的全体组成，不需要对组成集合元素的事和物的性质加以限制。常用的数的集合有

$N=\{1,2,\cdots,n,\cdots\}$ 表示正整数集,也可表示为 $N=\{n\}_{n=1}^{\infty}$;

$R=\{x \mid x \in(-\infty,+\infty)\}$ 表示全体实数集;

$C=\{z=x+iy \mid x,y \in R, i=\sqrt{-1}\}$ 表示全体复数集。

另外，本书中用 Q 表示全体有理数集，R^+ 表示全体正实数集，Z 表示全体整数集。在今后的讨论中会涉及到数域，这里先给出数域的概念。

数域：如果 F 是一个包含整数 0 和 1 的数集，且其中任两个数的和、差、积、商（除数不为零）仍是 F 中的数，则称 F 是一个数域。

讨论一个问题时指定数域就是规定了讨论的问题所涉及数的范围，这是必要的。上述数集中，实数集 R、复数集 C、有理数集 Q 都是数域，分别称为实数域、复数域和有理数域。又如集合 $\{a+b\sqrt{2} \mid a,\ b \in Q\}$ 也是一个数域。但正实数集 R^+ 和全体整数集 Z 都不是数域。

在讨论集合之间的关系以及逻辑命题的数学表述时，常用到如下的逻辑符号。

$\forall$：表示“任给一个”,“对每一个”(英文 Any 第一个字母倒写);

∃：表示“存在”（英文 Exist 第一个字母反写）；

⇒：推断符号。如 A 和 B 是两个命题，则“$A\Rightarrow B$”表示由 A 可推断出 B，但不能由非 A 否定 B，A 是 B 的充分条件。另外，“$A\Rightarrow B$”也表示 B 是 A 的必要条件，由非 B 可以否定 A，但不能由 B 推断出 A；

⇔：等价符号。如“$A\Leftrightarrow B$”表示 A 和 B 互为充分必要条件，即有“$A\Rightarrow B$”与“$B\Rightarrow A$”同时成立。

用逻辑符号可以简明而准确地表述数学中的逻辑命题（概念，定义、定理等）。一个正确的逻辑命题 P 可由 ∀ 开头（或 ∃ 开头），用若干个逻辑符号及随后叙述的内容相连，最后是命题的结论 Q 来表述，其形式为：命题 P：∀…；∃…；…；Q（命题的结论）。

而相反的命题非 P 可表述为：非 P：∃…；∀…；…；非 Q。

如“数列 $\{x_n\}$ 收敛于极限 a”可表述为

$$\forall \varepsilon > 0; \exists N_0 \in N; \forall n > N_0; 有 \mid x_n - a \mid < \varepsilon。$$

而相反的命题“数列 $\{x_n\}$ 不收敛于 a”可表述为

$$\exists \varepsilon > 0; \forall N_0 \in N; \exists n > N_0; 有 \mid x_n - a \mid > \varepsilon。$$

设 A 和 B 是两个集合，关于集合之间的关系及运算有以下一些重要的概念。

（1）子集与真子集

若 $\forall x \in A$ 有 $x \in B$，称 A 是 B 的子集，记为 $A \subset B$，也称 A 包含于 B。若 $A \subset B$ 和 $B \subset A$ 同时成立，则称 A 和 B 相等，记为 $A = B$。若 $A \subset B$，但 $A \neq B$ 且 $A \neq \Phi$，则 A 是 B 的真子集。

（2）余集

若 X 是在讨论问题范围中的全集，A⊂X，则记集合　$A^c = \{x \mid x \in X, x \overline{\in} A\}$，称为集合 A 的余集。

（3）集合的运算

集合之间的运算有并、交、差，分别记为“∪”、“∩”和“－”。定义为

$$A \cup B = \{x \mid x \in A 或 x \in B\}$$

$$A \cap B = \{x \mid x \in A 且 x \in B\}$$

$$A - B = \{x \mid x \in A 且 x \overline{\in} B\}$$

（4）集合的乘积

集合 A 与 B 的乘积定义为以下元素对的集合

$$A \times B = \{(x, y) \mid x \in A, y \in B\}$$

称为笛卡尔乘积。由此，平面上的点集 $P(x, y)$ 可用两个实数轴上的实数坐标对的集合表示为 $R^2 = R \times R = \{(x, y) \mid x \in R, y \in R\}$。进一步推广到全体实 n 维向量的集合可表示为 $R^n = R \times R \times \cdots \times R = \{(x_1, x_2, \cdots, x_n) \mid x_i \in R, i = 1, 2, \cdots, n\}$。

（5）映射

集合 A 到集合 B 的一个映射是指一个确定的法则 f，它使得 $\forall x \in A$，存在唯一的 $y \in B$，记为 $y = f(x)$。此时，y 称为 x 在映射 f 下的像，而 x 称为 y 在映射 f 下的一个原像。集合 A 称为 f 的定义域，而集合 $\{(x, y) = f(x) \mid x \in A\}$ 称为映射 f 的图形。映射 f 也常表示为

$$f: A \rightarrow B，即表示\ x \in A \rightarrow y = f(x) \in B$$

定义 设映射 $f:A \to B$

①若 $\forall x_1,x_2 \in A$,且 $x_1 \neq x_2$,都有 $f(x_1) \neq f(x_2)$,称 f 是一个单射;

②若 $\forall y \in B$,$\exists x \in A$,使得 $y = f(x)$,称 f 是一个满射;

③若 f 既是单射又是满射,则称它是双射或称为一一对应的映射。

如果映射 $f:A \to A$,它使每一个元素都对应到它自身。即 $\forall x \in A$,有 $f(x) = x$,则称 f 为一个恒等映射或单位映射,常记为 I_A。也可表示为 $\forall x \in A$,有 $I_A(x) = x$。恒等映射一定是一一对应的映射。

例 1-1 有映射 $f:R \to R, f(x) = x^2; g:R \to R^+, g(x) = x^2; h:R^+ \to R^+, h(x) = x^2$。以上三个映射虽由同一法则 $x \to x^2$ 定义,但 f 不是单射也不是满射;g 是满射,但不是一一对应的映射;h 是一一对应的映射,但不是恒等映射。

定义 设 $g:A \to B, f:B \to C$,若 $\forall x \in A$,对 x 相继施行映射g 和f,得到一个 $A \to C$ 的映射 $f \cdot g(x) = f(g(x))$,称 $f \cdot g:A \to C$ 是映射f 与g 的复合映射。复合映射也常简记为 fg,即 $\forall x \in A$,有 $fg(x) = f(g(x))$。

定义 设 $f:A \to B$。若存在映射 $g:B \to A$,使得有 $fg = I_B$ 和$gf = I_A$,称映射 f 是可逆映射,g 是f 的一个逆映射,常用 f^{-1} 表示 f 的逆映射,即 f^{-1} 满足 $ff^{-1} = I_B$ 和$f^{-1}f = I_A$。

映射 f 可逆的充要条件是f 是一一对应的映射。若映射 f 可逆,其逆映射是唯一的。

二、几种特殊的矩阵

矩阵是实数或复数按长方形排列的数组,一般 $m \times n$ 矩阵是由$m \times n$ 个数$a_{ij}(i = 1,2,\cdots,m;j = 1,2,\cdots,n)$,排成 m 行n 列的数表,记为 A

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \text{或 } A = (a_{ij})_{m\times n}$$

$n \times n$ 的方阵称为n 阶方阵。$a_{ij} \in R$ 的矩阵是实矩阵,$a_{ij} \in C$ 的矩阵是复矩阵,分别记为 $A \in R^{m\times n}$ 和$A \in C^{m\times n}$。

一个 $m \times n$ 矩阵A 可以按行向量分块

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \beta_1^T \\ \beta_2^T \\ \vdots \\ \beta_m^T \end{bmatrix}$$

其中 $\beta_i = (a_{i1},a_{i2},\cdots,a_{in})^T, i = 1,2,\cdots,m$。也可按列向量分块

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [\alpha_1,\alpha_2,\cdots,\alpha_n]$$

其中 $\alpha_j = (a_{1j},a_{2j},\cdots,a_{mj})^T, j = 1,2,\cdots,n$。

矩阵的秩:A 的行秩是A 的行向量组中最大线性无关组所含向量个数,A 的列秩是A 的列向量组中最大线性无关组所含向量个数。线性代数中已证明这两个数是相等的,这就是矩阵 A 的秩,记为 $r(A)$。矩阵 A 的秩$r(A)$定量地刻划了矩阵行(列)向量之间相互独立、线性无关的

程度,是反映矩阵 A 的本质属性的一个重要的数学量。求矩阵 A 的秩可以用矩阵的初等行(列)变换将 A 化为阶梯形矩阵,其不全为零的行(列)的个数就是矩阵的秩。由此也知,矩阵的初等变换不改变矩阵的秩。

如果 $A \in R^{m\times n}$,而 $r(A) = m \leqslant n$,称 A 为行满秩矩阵;若 $r(A) = n \leqslant m$,则称 A 为列满秩矩阵;若 $r(A) = m = n$,A 是非奇异的方阵, 也称为满秩方阵。

例 1-2

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 0 & -3 \\ 0 & 1 & 2 & 3 & 6 \\ 5 & 4 & 3 & 2 & -1 \end{bmatrix} \xrightarrow{\text{初等行变换}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad r(A) = 2$$

现介绍以下几种特殊形式和有特殊性质的矩阵。

1. 矩阵的转置和共轭转置阵

$A \in R^{m\times n}$,转置矩阵 $A^T \in R^{n\times m}$,$A^T = (a_{ji})_{n\times m}$。

$A \in C^{m\times n}$,共轭转置矩阵 $A^H \in C^{n\times m}$,$A^H = (\bar{a}_{ji})_{n\times m}$($\bar{a}_{ji}$ 是 a_{ji} 的共轭复数)。

2. 对角阵 $D \in R^{n\times n}$

$$D = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix} = \mathrm{diag}(d_1, d_2, \cdots, d_n)$$

对角阵有性质: $D^T = D, D^{-1} = \mathrm{diag}(\frac{1}{d_1}, \frac{1}{d_2}, \cdots, \frac{1}{d_n})(d_i \neq 0, i = 1,2,\cdots,n)$,$\det(D) = \prod_{i=1}^{n} d_i$,$D$ 的特征值$\lambda_i = d_i(i = 1,2,\cdots,n)$,当 $d_i = 1(i = 1,2,\cdots,n)$ 就是 n 阶单位阵$I_n = \mathrm{diag}(1,1,\cdots,1)$,常表示为 I。

3. 准对角阵 $A \in R^{n\times n}$

$$A = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_S \end{bmatrix}, \quad A^T = \begin{bmatrix} A_1^T & & & \\ & A_2^T & & \\ & & \ddots & \\ & & & A_S^T \end{bmatrix}$$

子块 $A_i \in R^{n_i\times n_i}(i = 1,2,\cdots,S)$,$\sum_{i=1}^{S} n_i = n$。若 A_i 可逆($i = 1,2,\cdots,S$),则 A 可逆,且有 $A^{-1} = \mathrm{diag}(A_1^{-1}, A_2^{-1}, \cdots, A_S^{-1})$,$\det(A) = \prod_{i=1}^{S}\det(A_i)$。

4. 三角阵 $A \in R^{n\times n}$

当 $i > j$ 时,$a_{ij} = 0$,A 为上三角阵;

当 $i < j$ 时,$a_{ij} = 0$,A 为下三角阵。

若 A 是上(下)三角阵,则 A^{-1} 也是上(下)三角阵(A 可逆),A^T 是下(上)三角阵,且 $\det(A) = \prod_{i=1}^{n} a_{ii}$,$A$ 的特征值$\lambda_i = a_{ii}(i = 1,2,\cdots,n)$。若 B 是与A 同阶的上(下)三角阵,则 AB 也是上

（下）三角阵。

5. 上 Hessenberg（海森伯格）阵 $A \in R^{n\times n}$

当 $i > j+1$ 时，$a_{ij}=0$，即次对角线以下均为零元，称 A 为上 Hessenberg 阵。

$$A=\begin{bmatrix} \times & \cdots & \cdots & \cdots & \times \\ \times & \times & \cdots & \cdots & \times \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \times \\ & & & \times & \times \end{bmatrix}$$

类似地还有下 Hessenberg 阵。一般上（下）Hessenberg 阵不考虑次对角线上有零元。

6.Hermite（埃尔米特）阵

$A \in C^{n\times n}$，若 $A^H = A$，称 A 为 Hermite 阵。如 $A=\begin{bmatrix} 3 & 2-i \\ 2+i & 4 \end{bmatrix} = A^H = \begin{bmatrix} 3 & \overline{2+i} \\ \overline{2-i} & 4 \end{bmatrix}$。

$A \in R^{n\times n}$，若 $A^T = A$，A 为实对称阵；若 $A^T = -A$，A 为实反对称阵。

7. 对角占优阵 $A \in R^{n\times n}$

若 $|a_{ii}| \geqslant \sum\limits_{\substack{j=1 \\ (j\neq i)}}^{n} |a_{ij}| \quad (i=1,2,\cdots,n)$，$A$ 为对角占优阵；

若 $|a_{ii}| > \sum\limits_{\substack{j=1 \\ (j\neq i)}}^{n} |a_{ij}| \quad (i=1,2,\cdots,n)$，$A$ 为严格对角占优阵。

可以证明，严格对角占优阵是非奇异阵。用反证法，先设 A 是严格对角占优且是奇异阵，由齐次线性方程组 $Ax=0$ 存在非零解可推出矛盾得证。

8. 正交阵，酉矩阵

设 $A \in R^{n\times n}$，若 $A^TA = AA^T = I$，即有 $A^T = A^{-1}$，称 A 为正交阵。正交阵的名称只用于实方阵。正交阵有以下性质：

① A 是正交阵，则 A 非奇异 $\det(A) = \pm 1$；

②A 是正交阵，则 $A^T = A^{-1}$ 也都是正交阵，且同级正交阵的乘积仍是正交阵；

③A 是正交阵的充分必要条件是它的行(列)向量组是法正交的向量组，即 A 的行(列)向量是互相正交的单位向量。

由正交条件 $AA^T = I$ 可得到 A 的行向量组法正交性质

$$\sum_{k=1}^{n} a_{ik}a_{jk} = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i\neq j \end{cases} \quad (i,j=1,2,\cdots,n) \tag{1-1}$$

由正交条件 $A^TA = I$ 可得到 A 的列向量组法正交性质

$$\sum_{k=1}^{n} a_{ki}a_{kj} = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i\neq j \end{cases} \quad (i,j=1,2,\cdots,n) \tag{1-2}$$

设 $A \in C^{n\times n}$，若 $A^HA = AA^H = I$，即有 $A^H = A^{-1}$，称 A 是酉矩阵。酉矩阵的名称是用于 n 阶复矩阵。同正交阵一样，对酉矩阵也可推出

$$\sum_{k=1}^{n} a_{ik}\bar{a}_{jk} = \delta_{ij} \text{ 和} \sum_{k=1}^{n} \bar{a}_{ki}a_{kj} = \delta_{ij} \qquad (i,j = 1,2,\cdots,n) \tag{1-3}$$

例 1-3 已知

$$A_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad A_3 = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

A_1 是正交阵。A_2 不是正交阵,它的行(列)向量虽然正交,但不是单位向量。A_3 也不是正交阵,因为正交阵必须是方阵,A_3 是列法正交的矩阵。

9. 正规矩阵

若 n 阶方阵 $A \in C^{n\times n}$(也包括 $A \in R^{n\times n}$)满足 $A^HA = AA^H$,则称 A 是正规矩阵。不难验证以上的对角阵、Hermite 阵、实对称阵、实反对称阵、正交阵和酉矩阵都是正规矩阵。当然还有其它满足条件的正规矩阵。如例 1-3 中 A_2 虽不是正交阵,但它却是一个正规矩阵。正规矩阵是一类重要的矩阵,它有一个重要的性质,即一个 n 阶方阵若是正规矩阵,其充分必要条件是它能用相似变换化简为一个对角阵。

10. 正定阵

正定阵是正定的实对称阵的简称,它是由正定的实二次型引入的。

实二次型:一个 n 元实二次型是 n 个实变量 $x_1,x_2,\cdots,x_n$ 的二次函数

$$\begin{aligned} f(x_1,x_2,\cdots,x_n) = {} & a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\ & + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n + \cdots \\ & + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 = \sum_{i,j=1}^{n} a_{ij}x_ix_j \end{aligned}$$

其中 $a_{ij} = a_{ji}(i,j = 1,2,\cdots,n)$。实二次型可用矩阵乘积形式表示为

$$f(x_1,x_2,\cdots,x_n) = x^TAx$$

其中 $x = (x_1,x_2,\cdots,x_n)^T \in R^n$,$A = (a_{ij})_{n\times n}$ 是一个实对称阵。

例 1-4 $f(x_1,x_2,x_3) = x_1x_2 + x_2x_3 + x_1x_3 + x_3^2$ 可表示为矩阵形式 $f(x_1,x_2,x_3) = x^TAx$。其中

$$A = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}, \qquad x = (x_1,x_2,x_3)^T$$

一个实二次型 $f(x_1,x_2,\cdots,x_n) = x^TAx$ 若对任意一组不全为零的实数 $c_1,c_2,\cdots,c_n$ 都有 $f(c_1,c_2,\cdots,c_n) > 0$,则称此实二次型是正定的。若"$>$"换成"$\geqslant$"0,则称此二次型是半正定的。

定义 设 A 为实对称阵，如果实二次型 x^TAx 是正定的，则 A 称为正定矩阵。

正定阵有如下重要性质：

① n 阶实对称阵是正定的⇔矩阵 A 的顺序主子式

$$\det(A_k)=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}>0 \quad (k=1,2,\cdots,n)$$

当 $k=n$ 时，有 $\det(A)=\det(A_n)>0$，因此 A 是非奇异的；或⇔矩阵 A 的特征值全大于零，即 $\lambda_i>0(i=1,2,\cdots,n)$；

②正定阵主对角元恒正。依次取 $x=e_i=(0,\cdots,0,\underset{(i)}{1},0,\cdots,0)^T\in R^n(i=1,2,\cdots,n)$，$x\neq0$，则由 A 是正定阵可得到 $a_{ii}=e_i^TAe_i>0(i=1,2,\cdots,n)$。

③ A 是正定阵，则 A^{-1} 也是正定阵；

④ $\forall A\in R^{n\times n}$，若 A 是非奇异的，则 A^TA 是 n 级实对称正定阵。

证明： A^TA 是对称阵显然成立。又因为 A 非奇异，必然列线性无关。因此，任给一非零向量 $x=(x_1,x_2,\cdots,x_n)^T\neq0$，必有 $Ax=y\neq0$。从而有

$$x^T(A^TA)x=(Ax)^TAx=y^Ty=\sum_{i=1}^{n}y_i^2>0$$

即 A^TA 是对称正定阵。此结论可推广到对 $\forall A\in R^{m\times n}$，若 A 列线性无关，则 A^TA 是 n 级对称正定阵。若 A 不能保证非奇异($A\in R^{n\times n}$)或不能保证列线性无关($A\in R^{m\times n}$)，则 A^TA 只能是半正定的。此时 n 级方阵的 $\lambda_i\geqslant0(i=1,2,\cdots,n)$。

三、初等矩阵

对矩阵的初等行（列）变换可以用初等矩阵与矩阵相乘表示出来。

定义 由单位矩阵经过一次初等变换得到的矩阵称为初等矩阵。

矩阵的初等变换有倍乘变换、对换变换和消元变换（把矩阵的某一行（列）倍乘后加到另一行（列）上）三种。每种初等变换都有一个初等矩阵与其对应。

1. 倍乘阵 $P(i(c))$

将单位阵 I 的第 i 行乘以非零常数 c。

$$\begin{array}{cl} P(i(c))=\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & c & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} & \begin{matrix} \\ \\ \\ \text{第 } i \text{ 行} \\ \\ \\ \\ \end{matrix} \end{array}$$

2. 对换阵 $P(i,j)$

将单位阵 I 的第 i 行和第 j 行互换。

$$P(i,j)=\left[\begin{array}{ccccccccccc} 1 &&&&&&&&&& \\ & \ddots &&&&&&&&& \\ && 1 &&&&&&&& \\ &&& 0 & \cdots & \cdots & \cdots & 1 &&& \\ &&& \vdots & 1 &&& \vdots &&& \\ &&& \vdots && \ddots && \vdots &&& \\ &&& \vdots &&& 1 & \vdots &&& \\ &&& 1 & \cdots & \cdots & \cdots & 0 &&& \\ &&&&&&&& 1 && \\ &&&&&&&&& \ddots & \\ &&&&&&&&&& 1 \end{array}\right]\begin{array}{l} \\ \\ \\ 第\ i\ 行 \\ \\ \\ \\ 第\ j\ 行 \\ \\ \\ \\ \end{array}$$

3. 消元阵 $P(i,\ j\ (k))$

将单位阵 I 的第 j 行的 k 倍加到第 i 行。

$$P(i,j(k))=\left[\begin{array}{cccccccc} 1 &&&&&&& \\ & \ddots &&&&&& \\ && 1 & \cdots & \cdots & k && \\ &&& \ddots && \vdots && \\ &&&& \ddots & \vdots && \\ &&&&& 1 && \\ &&&&&& \ddots & \\ &&&&&&& 1 \end{array}\right]\begin{array}{l} \\ \\ 第\ i\ 行 \\ \\ \\ 第\ j\ 行 \\ \\ \\ \end{array}$$

以上三种初等矩阵都是非奇异的，且它们的逆矩阵仍然是初等矩阵。不难验证

$$(P(i,j))^{-1}=P(i,j)$$

$$(P(i(c)))^{-1}=P(i(\frac{1}{c}))$$

$$(P(i,j(k)))^{-1}=P(i,j(-k))$$

矩阵的初等变换可用初等矩阵与矩阵相乘表示。

$\forall A\in R^{m\times n}$

对 A 施行一次初等行变换相当于对矩阵 A 左乘 m 阶初等矩阵

对 A 施行一次初等列变换相当于对矩阵 A 右乘 n 阶初等矩阵

下面再介绍一种常用的初等矩阵——高斯（Gauss）变换阵。

4. Gauss 变换阵 L_j

设向量 $\bar{l}_j=(0,\cdots,0,l_{j+1,j},l_{j+2,j},\cdots,l_{n,j})^T\in R^n$ 和 $e_j=(0,\cdots,0,\underset{(j)}{1},0,\cdots,0)^T\in R^n$，显然有 $e_k^T\bar{l}_j=0(j=1,2,\cdots,n-1;k\leqslant j)$。定义 Gauss 变换阵为 $L_j=I-\bar{l}_je_j^T$，用矩阵乘法展开后得到