【美】杰洛斯拉夫·达奇纳
王统　张尧弼译　李玉兴校

数值分析及其程序应用手册

上海科学技术文献出版社

数值分析及其程序应用手册

〔美〕 杰洛斯拉夫·达奇纳

王 统 张尧弼 译

李玉兴 校

上海科学技术文献出版社

数值分析及其程序应用手册

［美］ 杰洛斯拉夫·达奇纳

王 统 张尧弼 译

李玉兴 校

*

上海科学技术文献出版社出版发行

（上海市武康路2号）

新 华 书 店 经 销

昆山亭林印刷厂印刷

*

开本 787×1092 1/16 印张 20 字数 499,000

1988年12月第1版 1988年12月第1次印刷

印数：1—3,000

ISBN 7-80513-292-5/O·25

定价 12.20元

《科技新书目》181-224

译　序

工程技术和科学研究中经常遇到难以用解析法求解的数学问题，它们可以用数值法求解。电子计算机特别是微型机的广泛应用，使数值分析方法得到了飞速发展，在科研、工程技术和教学各领域起着极为重要的作用。

这本手册为科学家、工程师、高校师生和广大科技工作者提供了简捷、精确的数值解法的研究结论和非常实用计算程序，其目的在于使它成为人们手中犀利的分析工具。手册在数值分析、微机应用和广大读者、用户之间架起了相互沟通的桥梁，因此，对我国广大科技工作者十分适用。

手册内容共分为两大部分：第一部分简明扼要地介绍了数值分析方法的主要原理和结论；第二部分为实用计算程序，是作者多年来在高等院校教学和研究中积累的经验结晶，都是极为珍贵的资料。手册中的88个计算程序均用BASIC语言写成，可以很方便地在普通微机上使用。在编辑过程中，为节省篇幅，略去了第二部分中的例题。我们相信，在使用这本手册后，读者和用户将能获得无可比拟的财富——研究人员的宝贵时间。

本书的译校分工如下：序言、程序使用说明及程序中的文字部分由王统、张尧弼同志合译，第一、第二、第三、第四章由张尧弼同志翻译，第五、第六、第七、第八、第九章由王统同志翻译。全书由李玉兴、张培伟等同志总校，在文稿整理和校对工作中得到张培伟同志的大力帮助，谨此致谢！

限于我们的水平，中译本中错误和不当之处在所难免，恳请广大读者和用户提出宝贵意见。

译　者

一九八八年八月二十日

目　　录

第一部分　理　论

第二部分 BASIC 程序

原　　序

随着计算机、微型计算机和可编程计算器等计算工具的出现，过去许多在工程上和科学研究上有着重要应用，但又难以用解析方法来求解的数学问题，现在都可以当作数值计算问题来处理了。通常，在一些数值分析的书中，在介绍用某种方法求解一个给定的问题时，总是或多或少地对其进行繁琐的推导和证明，对于它的应用范围进行详尽周密的考虑和说明。当然，这作为初学者的基本训练是绝对必要的。但是，对于多数从事研究工作的工程师、科学家或者研究生来说，他们所需要的则是简捷而精确的，可用于求解问题的某些数值计算方法的结论，如有可能，还应尽量提供对问题进行数值计算处理的相应的计算机程序。本书的目的，正是为了提供这方面的有关资料，以便节省研究人员的宝贵时间。

本书分为两个部分。第一部分包括九章，每章开头都有一个简短的有关分析背景的说明，对最有用的一些数值方法以及在应用中所必须了解的问题都作了必要的介绍，但不作详细推导和证明，因为读者可在有关的著作中查到它们。对于所介绍的数值方法，重点放在讨论数值计算的稳定性、精度和速度方面。通常，研究人员可在几个并列的方法之间进行选择。当某一种方法明显优于其他方法时，将只讨论这一种方法。第二部分是八十五个程序。在附录中还有三个附加的程序。读者可根据自己的需要从中进行挑选和使用。

第一章论述了线性代数的内容：线性代数方程组、矩阵代数、行列式求值和超定方程组。所讨论的数值解法包括直接法、选取部分主元的和行列式求值的高斯消元法及其对它的修正（三对角线和五对角线矩阵）和迭代法。

第二章涉及到这样一个重要问题，即用一个解析表达式或一组数字数据定义的函数的插值和逼近问题，这里说的数字数据可以认为是精确的（如由数字表格提供的），也可以认为它包含着一些随机的误差（如在前期的数字计算中观测或产生的）。该函数或者用一个多项式来插值，或者由几个正交多项式级数或函数（最小二乘法）来逼近，或者由一个有理函数来逼近。插值和逼近的思想不仅适用于函数，而且也适用于它的一阶和二阶导数。此外本章还介绍了埃尔米特插值、三样条和微分公式等。

第三章论述了在有限、半有限、无限区间上积分的数值计算问题和被积函数由一个解析表达式或由一组等距或不等距的数字数据定义的数值积分问题。此外，奇异积分的求值和相应的数值方法也在这一章中进行了讨论。

第四章中将讨论对常微分方程组积分的解析方法和数值方法。在对欧拉和泰勒级数方法进行简短的叙述之后，还对龙格-库塔法和预测-校正法进行了广泛的讨论，其中包括只与一个初值有关的新的而且有用的预测-校正法。

常微分方程组的边值问题将在第五章中（从一个可见的解析点）讨论，在那里将用一种数值方法把边值问题变成第四章中的初值问题来处理。

第六章从对二次、三次和双二次多项式求根公式和程序的分析开始，来讨论非线性代数方程和超越方程的求解问题，以及用迭代法计算一个代数方程或超越方程的实根，或者计算一个多项式的实根和复根的问题。

第七章仍回到对线性代数的讨论，它介绍了特征值和特征向量的计算问题。第八章包括了定义特殊函数的公式和求解它们的程序。第九章讨论了数理统计方法的选择、曲线拟合以及离散与连续分布的问题。

附录中第一个程序用来以单精度计算已得出的十三个初等函数。第二个程序用来以双精度计算十八个初等函数，当这些函数不包含在微型计算机的标准函数库中时，可使用这些程序。这时，当确信程序是采用双精度运行时，第三个程序的一段或几段必须作为子程序调用。

由于这些程序都考虑用于微型计算机，所以必须限制所处理的是实矩阵问题或实变量问题。虽然实系数的非线性代数方程组的根可能是复共轭的，但这只是一个仅有的例外。所以，尽管一个代数问题的特征值可能是复共轭的（因为假定为实矩阵），但特征向量的计算还必须限定为简单的实型特征值矩阵。

当使用计算机时，选择编程语言是很重要的。用 FORTRAN 语言编写的源程序读起来比用 BASIC 语言编写的更清楚。但另一方面，用 FORTRAN 语言编写的程序在编译和连接时需要较多的内存空间，而且，当改变数组的维数或调整程序以便处理另一个问题时，例如不同的被积函数或不同的微分方程组等，则程序必须重新编译和连接。而用 BASIC 语言编程，则只需把源程序保存在内存中即可，因为源程序的编译是自动的，而且，在进行编译时，快得连使用者都感觉不到。此外，对 BASIC 语言来说，数组的维数是在每次执行时由输入的数据来设定的，而且并不需要按照一定的格式（在 FORTRAN 中格式化的输出数据被认为是它的特点，而并不认为是一种不足）。然而，具有决定意义的是计算速度，尽管，在同样的微型计算机上运行用这两种语言编写的同一程序时，BASIC 程序将增加 10～20% 的计算时间，但这并不妨碍在本书中选用 BASIC 语言来编程，因为它的优点无疑胜过它的不足。

研究人员手头的微型计算机是用来求解本书中所有问题的数值解的理想工具，因为它的计算速度已足够快。通常，完成一个计算任务的时间将不会超过一分钟，最多不超过几分钟。当然，以高精度计算一组常微分方程组的积分，如程序 P406 的例子，可能要化一个多小时，然而这在大型计算机中计算则只需几秒钟即可完成。但是这样的比较是很容易令人产生误解的，因为对微型计算机来说，程序一旦递交运行，则运行立即开始；而对于大型计算机来说，通常不仅必须等待程序进入计算机运行，而且还要等待结果的打印输出。此外，使用大计算机的费用可能是相当高的。作者相信，今后推出的速度更快、内存更大的微型计算机将能解决在科学研究中所遇到的所有问题，甚至包括二阶偏微分方程组的积分。

本书是作者在讲授超前分析和数值分析时的经验总结，是基于作者的笔记和写了几年的计算程序的汇集。

本书的公开出版受到了 McGraw-Hill 图书公司的 Potricia Allen-Browne 通力协作和 Elizabeth. P. Richardson 对原稿的编辑并在使其付印方面的协助。在结束前言之前，请允许作者对他们表示诚挚的谢意。

第一章 线性代数I

线性代数中有两个基本问题：线性代数方程组的求解和求矩阵的特征值。前者虽然可通过有限次的四种基本算术运算精确地求解，但是当这些运算的数量太大，以致不得不考虑舍入误差的影响时，或者所讨论的方程组为病态时，求解是很困难的。本章将论述如何解决这些困难的方法。关于特征值问题，可以使用某种迭代法来求解，或者把它简化为一个特征多项式的求解问题。由于非线性方程的求解问题要在第六章中论述，所以特征值问题就移到第七章中去讨论了。

§1.1 线性代数方程组

讨论一个含有 n 个未知量 m个方程的线性代数方程组

$$\begin{aligned}
&a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1\\
&a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&a_{m1}x_1+a_{m2}x+\cdots+a_{mn}x_n=b_m
\end{aligned}\tag{1.1.1a}$$

写成矩阵形式是

$$AX=B\tag{1.1.1b}$$

本章中用大写字母表示矩阵，其中 A 就是一个 m 行 n 列的长方形矩阵

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{bmatrix}\tag{1.1.2}$$

X 和 B 是单列矩阵

$$X=\begin{bmatrix} x_1\\ x_2\\ \vdots\\ x_m\end{bmatrix}\qquad B=\begin{bmatrix} b_1\\ b_2\\ \vdots\\ b_m\end{bmatrix}\tag{1.1.3}$$

如果方程的个数等于未知量的个数，即 $m=n$，则方程组是确定的。而如果确定的方程组中每个方程与其它方程是线性无关的，则矩阵 A 有逆矩阵 A^{-1}。

其定义为

$$A^{-1}A=AA^{-1}=I\tag{1.1.4}$$

其中 I 是单位矩阵，即非零元素都等于1的对角矩阵

$$I=\begin{bmatrix} 1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & 1\end{bmatrix}\tag{1.1.5}$$

根据式(1.1.4)，方程(1.1.1b)可左乘一个A^{-1}，未知量x_i便可精确地确定

$$X=A^{-1}B \tag{1.1.6}$$

但随着未知量个数的增加，逆矩阵的计算也将越来越复杂。在下面几节中，将讨论解方程组(1.1.1a)的更有效的方法。

如果在式(1.1.1a)中，方程数少于未知量的个数，即$m<n$，则方程组是不确定的。在这种情况下，$n-m$个未知量x_i可以自由选择，从而简化为一个确定的方程组，这些方程将是相互独立的。

如果在式(1.1.1a)中方程数多于未知量的个数，即$m>n$，则方程组是超定方程组，这在测量次数大于未知量个数(以减小观测误差的影响)时是经常发生的。超定方程组不能精确地求解，最理想的求解方法是用最小二乘法(见§1.6)。

§1.2 矩 阵

矩阵A是一个具有$m\times n$个元素a_{ij}的数组，排列成m行n列，如式(1.1.2)所示。如果$m=1$，则矩阵为单行矩阵，即行向量；如果$n=1$，则为单列矩阵，如式(1.1.3)所示，称为列向量。如果在缺少元素的位置上填上零，那么每个矩阵都可以写成一个$n\times n$的方阵($m=n$)，并称为n阶方阵。一个所有元素都为零的矩阵称为零矩阵或空矩阵，记作“O”。

除主对角线上的元素之外，其余所有元素都为零的矩阵称作对角矩阵，即

$$d_{ii}\neq 0 \qquad d_{ij}=0 \quad \text{当} \quad i\neq j \tag{1.2.1}$$

对角矩阵的特殊情况是单位矩阵I，或称么阵，用δ表示[参看式(1.1.5)]

$$\delta_{ij}=\begin{cases}1 & \text{当} \quad i=j\\ 0 & \text{当} \quad i\neq j\end{cases} \tag{1.2.2}$$

三对角线矩阵T，其沿主对角线方向窄带状区域内的元素为非零

$$t_{ii}\neq 0 \qquad t_{i,i\pm 1}\neq 0 \qquad \text{其余} \quad t_{ij}=0 \tag{1.2.3}$$

与此相类似，五对角线矩阵F是

$$f_{ii}\neq 0 \qquad f_{i,i\pm 1}\neq 0 \qquad f_{i,i\pm 2}\neq 0 \qquad \text{其余} \quad f_{ij}=0 \tag{1.2.4}$$

n行n列的下三角矩阵L，其主对线及其下方的元素均为非零，而主对角线上方的元素均为零。即

$$\begin{aligned} l_{i,i-k}&\neq 0 \qquad k=0,1,\cdots,i-1\\ l_{i,i+k}&=0 \qquad k=1,2,\cdots,n-1 \end{aligned} \tag{1.2.5}$$

同样，n行n列的上三角矩阵U，其主对角线及其上方的元素均为非零，而主对角线下方的元素均为零。即

$$\begin{aligned} u_{i,i+k}&\neq 0 \qquad k=0,1,\cdots,n-i\\ u_{i,i-k}&=0 \qquad k=1,2,\cdots,i-1 \end{aligned} \tag{1.2.6}$$

如果矩阵中几乎所有元素都为非零，则该矩阵称作满阵，而稀疏矩阵则是其中元素只有少数为非零，它们通常都在沿主对角线的窄带区域内。

转置矩阵$\tilde{A}$是由交换A矩阵的行与列而得。即

$$\tilde{a}_{ij}=a_{ji} \tag{1.2.7}$$

复共轭矩阵A^*是对矩阵A的每个元素取共轭复数后而得。即

$$a^*_{ij}=(a_{ij})^* \tag{1.2.8}$$

伴随矩阵 A^+ 是由转置一个复共轭矩阵 A^*，或对转置矩阵 $\tilde{A}$ 求其复共轭矩阵而产生

$$A^\dagger = \tilde{A}^* = \tilde{A}^* \tag{1.2.9}$$

借助于这些运算和由式 (1.1.4) 定义的逆阵运算，对具有独特性质的特殊方阵定义如下：

实矩阵 $$A^* = A \tag{1.2.10}$$

对称矩阵 $$\tilde{A} = A \tag{1.2.11}$$

斜对称矩阵(反对称矩阵) $$\tilde{A} = -A \tag{1.2.12}$$

正交矩阵 $$A^{-1} = \tilde{A} \tag{1.2.13}$$

纯虚数矩阵 $$A^* = -A \tag{1.2.14}$$

埃尔米特矩阵 $$A^\dagger = A \tag{1.2.15}$$

斜埃尔米特矩阵 $$A^\dagger = -A \tag{1.2.16}$$

酉矩阵 $$A^{-1} = A^\dagger \tag{1.2.17}$$

在式(1.2.15)～(1.2.17)中，如果矩阵 A 是实矩阵，则埃尔米特矩阵、斜埃尔米特矩阵和酉矩阵将分别蜕化为如式(1.2.11)～(1.2.13)所示的对称矩阵、斜对称矩阵和正交矩阵。

这些矩阵的特殊性质是由于它们的元素具有特殊对称的缘故，它减少了独立元素的个数，例如在 n 阶的对称矩阵中，只有 $n(n+1)/2$ 个元素是独立的。如果考虑矩阵元素的大小，则在求解一个线性代数方程组时，有两种特殊的矩阵是很重要的，其一是严格的主对角线优势矩阵，它由条件

$$|a_{ii}| \geqslant \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \quad i = 1, 2, \cdots, n \tag{1.2.18}$$

定义；其二是对称正定矩阵，即满足如下附加条件的对称矩阵：对任意向量 X，有

$$\tilde{X}AX > 0 \tag{1.2.19}$$

一个 n 阶对称矩阵是正定矩阵的充要条件是

$$\det A_k > 0 \qquad k = 1, 2, \cdots, n \tag{1.2.20a}$$

其中 A_k 是矩阵 A 的前 k 行和前 k 列相交而形成的 $k \times k$ 矩阵(西勒维斯特准则)，所以对角线上的元素都为正

$$a_{ii} > 0 \qquad i = 1, 2, \cdots, n \tag{1.2.20b}$$

并且

$$|a_{ij}|^2 \leqslant a_{ii} a_{jj} \qquad i, j = 1, 2, \cdots, n \tag{1.2.20c}$$

假定有矩阵 A 和矩阵 B，如果其对应的元素都相等，即

$$a_{ij} = b_{ij} \tag{1.2.21}$$

则称矩阵 A 和 B 相等。

矩阵 A 和矩阵 B 对应的元素相加或相减，可生成一个新的矩阵

$$C = A \pm B = \pm B + A$$

其元素是

$$c_{ij} = a_{ij} \pm b_{ij} \tag{1.2.22}$$

矩阵 A 乘以一个标量可生成一个新的矩阵 $B = gA$，其元素为该标量乘 A 中每个对应元素，即

$$b_{ij} = g a_{ij} \tag{1.2.23}$$

矩阵 A 和矩阵 B 相乘可生成一个新的矩阵 $C = AB$，其元素为

$$c_{ij} = \sum_k a_{ik} b_{kj} \tag{1.2.24}$$

因此，一般说来

$$AB \neq BA \qquad AI = IA = A \tag{1.2.25}$$

如果 $AB = BA$，则称矩阵A，B 是可交换的。

两矩阵的积 AB 的转置等于它们转置矩阵的积，但次序颠倒。

$$\widetilde{AB} = \widetilde{B}\widetilde{A} \tag{1.2.26}$$

方阵 A 的迹是其对角线元素的和，即

$$\mathrm{Tr}A = \sum_i a_{ii} \tag{1.2.27}$$

因此，

$$\mathrm{Tr}(AB) = \sum_i \sum_k a_{ik} b_{ki} = \sum_k \sum_i b_{ki} a_{ik} = \mathrm{Tr}(BA) \tag{1.2.28}$$

即使 $AB \neq BA$，上式也成立。

§1.3 行列式和逆矩阵

方阵是一个 $n \times n$ 的数组，通常与其 n 阶行列式的值有关。而行列式

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

由下式定义

$$\det A = \sum (-1)^S a_{i1} a_{j2} \cdots a_{pn} \tag{1.3.1b}$$

和式中包括所有的 $i \neq j \neq \cdots\cdots \neq n$ 的排列，其中 S 是序列(i, j, k, …, p)交换成自然序列(1, 2, 3, …, n)的交换次数。例如

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} a_{22} - a_{12} a_{21} \tag{1.3.2}$$

$$\det B = \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} = b_{11}b_{22}b_{33} - b_{11}b_{32}b_{23} + b_{31}b_{12}b_{23} - b_{31}b_{22}b_{13} + b_{21}b_{32}b_{13} - b_{21}b_{12}b_{33} \tag{1.3.3}$$

如果 n 阶行列式用式(1.3.1b)来求值，则其值由 $n!$ 项的代数和给定。由于通常它的值远小于代数和中某项的值，并且还必须用高精度来计算，以减小舍入误差的影响，所以，当行列式的阶数 $n > 3$ 时，一般很少用式(1.3.1b)来求值。

最好的办法是把矩阵 A 分解为下三角矩阵与上三角矩阵的积来求，即

$$A = LU \tag{1.3.4}$$

由于任意两个具有相同阶数的矩阵，可应用定理

$$\det(AB) = (\det A)(\det B) \tag{1.3.5}$$

根据式(1.3.1b)，下三角矩阵或上三角矩阵的行列式等于其对角线上各项的积

$$\det L = \prod_{i=1}^{n} l_{ii} \quad \det U = \prod_{i=1}^{n} u_{ii} \tag{1.3.6}$$

另外，在§1.4中将证明在LU矩阵的分解中，L矩阵的对角线项都等于1，因此

$$\det A=\prod_{i=1}^{n}u_{ii} \tag{1.3.7}$$

下述的几个定理在行列式求值中是很有用的：

• 交换行列式中任意两行(或两列)，则行列式将只改变符号。

• 如果一个行列式的所有行和所有列都互相交换，则行列式的值不变：

$$\det A=\det \tilde{A}。$$

• 如果一行列式中某行(或列)的所有元素同乘以一数，并加到另一行(或列)的各对应元素上，则行列式的值不变。

• 如果行列式中一行(或列)等于零，则行列式的值等于零。

• 作为后两个定理的推论，如果行列式中任意两行(或列)相同或互相成比例，则行列式的值为零。

• 行列式中某行(或列)的每个元素都乘以一个标量，则等于该行列式乘以这个标量。

定义：

• 行列式的值为0的矩阵叫奇异矩阵，反之为非奇异矩阵。

• 如果一个r阶非奇异矩阵A_r可由一个n阶的奇异矩阵A消去$n-r$行和列而得到，则矩阵A具有秩r，并且把非奇异矩阵A_r叫作A的基。很明显$r\leqslant\min(m, n)$，其中m, n分别是矩阵A的行数和列数。

• 方阵A的子式$\mathrm{Min}\ a_{ij}$是由该矩阵的行列式去掉第i行和第j列中所有元素后获得的。

• 方阵A的代数余子式$\mathrm{Cof}\ a_{ij}$等于子式$\mathrm{Min}\ a_{ij}$乘以$(-1)^{i+j}$，即

$$\mathrm{Cof}\ a_{ij}=(-1)^{i+j}\,\mathrm{Min}\ a_{ij} \tag{1.3.8}$$

根据这个定义，行列式可由拉普拉斯展开式来计算

$$\det A=\sum_{k=1}^{n} a_{ik}\,\mathrm{Cof}\ a_{ik}=\sum_{k=1}^{n} a_{ki}\,\mathrm{Cof}\ a_{ki} \tag{1.3.9}$$

其中i是任意指定的。根据这个公式，行列式的值等于第i行（或列）中的每个元素乘以相应的余子式的代数和。

由式(1.1.4)定义，逆阵A^{-1}中的元素为

$$a_{ij}^{-1}=\frac{\mathrm{Cof}\ a_{ji}}{\det A} \tag{1.3.10}$$

因此，只有非奇异矩阵才有其逆矩阵。

逆矩阵具有两个重要的性质

$$(AB)^{-1}=B^{-1}A^{-1} \quad 和 \quad (A^{-1})^{-1}=A \tag{1.3.11}$$

对于对角矩阵和下三角矩阵或上三角矩阵，一般公式(1.3.10)具有简单的形式

$$d_{ii}^{-1}=\frac{1}{d_{ii}} \quad 其余 \quad d_{ij}^{-1}=0 \tag{1.3.12}$$

$$l_{ii}^{-1}=\frac{1}{l_{ii}} \qquad l_{ij}^{-1}=\frac{\sum_{k=j}^{i-1} l_{ik}l_{kj}^{-1}}{l_{ii}} \qquad i=j+1, \cdots, n$$
$$其余 \quad l_{ij}^{-1}=0 \tag{1.3.13}$$

数值関