Basic Numerical Mathematics
ISNM

INTERNATIONAL SERIES OF NUMERICAL MATHEMATICS
INTERNATIONALE SCHRIFTENREIHE ZUR NUMERISCHEN MATHEMATIK
SÉRIE INTERNATIONALE D'ANALYSE NUMÉRIQUE

Editors:
Ch. Blanc, Lausanne; A. Ghizzetti, Roma; P. Henrici, Zürich; A. Ostrowski, Montagnola;
J. Todd, Pasadena; A. van Wijngaarden, Amsterdam

VOL 22
Basic Numerical Mathematics

Vol. 2: Numerical Algebra

by

John Todd
Professor of Mathematics
California Institute of Technology
Contents

Notations and Abbreviations ......................................................... 7
Preface ....................................................................................... 9
Chapter 1. Manipulation of Vectors and Matrices .................................. 13
Chapter 2. Norms of Vectors and Matrices ........................................... 16
Chapter 3. Induced Norms .............................................................. 19
Chapter 4. The Inversion Problem I: Theoretical Arithmetic .................... 29
Chapter 5. The Inversion Problem II: Practical Computation ................. 44
Chapter 6. The Characteristic Value Problem — Generalities ............... 53
Chapter 7. The Power Method, Deflation, Inverse Iteration .................. 65
Chapter 8. Characteristic Values .................................................... 71
Chapter 9. Iterative Methods for the Solution of Systems $A\mathbf{x} = b$ ........ 83
Chapter 10. Application: Solution of a Boundary Value Problem .......... 99
Chapter 11. Application: Least Squares Curve Fitting ......................... 105
Chapter 12. Singular Value Decomposition and Pseudo-Inverses ........... 110
Solutions to Selected Problems ..................................................... 117
Bibliographical Remarks ............................................................. 212
Index ......................................................................................... 214
Notations and Abbreviations

\(\delta(i, j)\) or \(\delta_{ij}\) is defined by: \(\delta(i, j) = 0\) if \(i \neq j\), \(\delta(i, i) = 1\).

\(\mathbb{R}_n\) (resp. \(\mathbb{C}_n\)) denotes the space of column vectors with \(n\) real (resp. complex) components.

Vectors will be denoted by lower case black type so that, e.g., if \(a \in \mathbb{R}_n\) then

\[
a = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}
\]

where the \(a\)'s are real.

\(e_i\) will denote the \(i\)th unit column vector, all components being zero except the \(i\)th, which is 1 so that \([e_i]_j = \delta(i, j)\).

\(e\) will denote the column vector all of whose components are 1.

Matrices will be denoted by upper case black type so that, e.g.,

\[
A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [a_{ij}]
\]

is an \(m \times n\) matrix. We shall usually dispense with commas between the elements \(a_{ij}\) and between their subscripts.

\(I\) or \(I_n\) will denote the unit \(n \times n\) matrix: \([I]_{ij} = \delta(i, j)\).

\(J\) will denote the matrix all of whose elements are 1.

If \(a_{ij} = 0\) for \(i \neq j\) we shall write \(A = \text{diag} \{a_{11}, \ldots, a_{nn}\}\).

Transposition of matrices (including vectors) will be denoted by a prime or dash so that \(A'\) will be an \(n \times m\) matrix and \(a'\) a row vector. We will use * to denote conjugate transposition so that, e.g.,

\[
[A^*]_{ij} = \bar{a}_{ji} \quad \text{and} \quad a^* = [\bar{a}_1, \ldots, \bar{a}_n].
\]

We shall occasionally write \(\text{row}_i\) to indicate the \(i\)th row of a matrix and similarly use \(\text{col}_j\).

The inner (or scalar) product of two vectors will be denoted by

\[
(x, y) = x'y = \sum x_i y_i.
\]

The determinant of a square matrix \(A\) will be denoted by \(\det A\).

The trace of a square matrix \(A\) is \(\text{tr} A = \sum a_{ii}\).
The characteristic polynomial of a square matrix $A$ is $\det(A - xI)$. The characteristic values of $A$ are the zeros of its characteristic polynomial and are usually denoted by $\lambda_1, \ldots, \lambda_n$ and it is often assumed that $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$. In the latter case $\lambda_1$ is called the dominant characteristic value and $|\lambda_1|$ the spectral radius of $A$.

A characteristic vector of $A$ corresponding to a characteristic value $\lambda$ is a (non-zero) solution of the equation $Aa = \lambda a$. We then speak of $(\lambda, a)$ as a characteristic pair of $A$.

The pseudo-inverse of a (rectangular) matrix $A$ will be denoted by $A^\dagger$; when $A$ is square and $\det A \neq 0$, $A^\dagger = A^{-1}$.

Norms of vectors and matrices will be denoted by $\|a\|$ or $\|A\|$ with appropriate subscripts.

We use the standard order symbolism, e.g., $f(n) = O(g(n))$ as $n \to \infty$ means that there exists $n_0, A$ such that when $n \geq n_0$ (depending in general on $A$) we have $|f(n)| \leq A|g(n)|$. Often $g(n)$ will be a power of $n$.

$\sum'$ will be used to indicate (or emphasize) the omission of a special term in a sum, e.g.,

$$\sum'|a_{ij}| = \sum_{j \neq i} a_{ij}.$$

We use $\doteq$ to indicate approximate equality.

We use the standard logical symbolism e.g., $\Rightarrow$ for implies, $\in$ for belongs to, $\notin$ for does not belong to, $\subset$ for is included in.
Preface

There is no doubt nowadays that numerical mathematics is an essential component of any educational program. It is probably more efficient to present such material after a reasonable competence in (at least) linear algebra and calculus has already been attained — but at this stage those not specializing in numerical mathematics are often interested in getting more deeply into their chosen field than in developing skills for later use. An alternative approach is to incorporate the numerical aspects of linear algebra and calculus as these subjects are being developed. Long experience has persuaded us that a third attack on this problem is the best and this is developed in the present two volumes, which are, however, easily adaptable to other circumstances.

The approach we prefer is to treat the numerical aspects separately, but after some theoretical background. This is often desirable because of the shortage of persons qualified to present the combined approach and also because the numerical approach provides an often welcome change which, however, in addition, can lead to better appreciation of the fundamental concepts. For instance, in a 6-quarter course in Calculus and Linear Algebra, the material in Volume 1 can be handled in the third quarter and that in Volume 2 in the fifth or sixth quarter.

The two volumes are independent and can be used in either order — the second requires a little more background in programming since the machine problems involve the use of arrays (vectors and matrices) while the first is mostly concerned with scalar computation.

In the first of these, subtitled "Numerical Analysis", we assume that the fundamental ideas of calculus of one variable have been absorbed: in particular, the ideas of convergence and continuity. We then take off with a study of "rate of convergence" and follow this with accounts of "acceleration process" and of "asymptotic series" — these permit illumination and consolidation of earlier concepts. After this we return to the more traditional topics of interpolation, quadrature and differential equations.

Throughout both volumes we emphasize the idea of "controlled computational experiments": we try to check our programs and get some idea of
errors by using them on problems of which we already know the solution — such experiments can in some way replace the error analyses which are not appropriate in beginning courses. We also try to exhibit “bad examples” which show some of the difficulties which are present in our subject and which can curb reckless use of equipment. In the Appendix we have included some relatively unfamiliar parts of the theory of Bessel functions which are used in the construction of some of our examples.

In the second volume, subtitled “Numerical Algebra”, we assume that the fundamental ideas of linear algebra: vector space, basis, matrix, determinant, characteristic values and vectors, have been absorbed. We use repeatedly the existence of an orthogonal matrix which diagonalizes a real symmetric matrix; we make considerable use of partitioned or block matrices, but we need the Jordan normal form only incidentally. After an initial chapter on the manipulation of vectors and matrices we study norms, especially induced norms. Then the direct solution of the inversion problem is taken up, first in the context of theoretical arithmetic (i.e., when round-off is disregarded) and then in the context of practical computation. Various methods of handling the characteristic value problems are then discussed. Next, several iterative methods for the solution of system of linear equations are examined. It is then feasible to discuss two applications: the first, the solution of a two-point boundary value problem, and the second, that of least squares curve fitting. This volume concludes with an account of the singular value decomposition and pseudo-inverses.

Here, as in Volume 1, the ideas of “controlled computational experiments” and “bad examples” are emphasized. There is, however, one marked difference between the two volumes. In the first, on the whole, the machine problems are to be done entirely by the students; in the second, they are expected to use the subroutines provided by the computing system — it is too much to expect a beginner to write efficient matrix programs; instead we encourage him to compare and evaluate the various library programs to which he has access.

The problems have been collected in connection with courses given over a period of almost 30 years beginning at King's College, London, in 1946 when only a few desk machines were available. Since then such machines as SEAC, various models of UNIVAC, Burroughs, and IBM equipment and, most recently, PDP 10, have been used in conjunction with the courses which have been given at New York University, and at the California Institute of Technology.

We recommend the use of systems with “remote consoles” because, for instance, on the one hand, the instantaneous detection of clerical slips and,
on the other, the sequential observation of convergents is especially valuable to beginners. The programming language used is immaterial. However, most of the problems in Volume 1 can be dealt with using simple programmable hand calculators but many of these in Volume 2 require the more sophisticated hand calculators (i.e. those with replaceable programs).

The machine problems have been chosen so that a beginning can be made with very little programming knowledge, and competence in the use of the various facilities available can be developed as the course proceeds. In view of the variety of computing systems available, it is not possible to deal with this aspect of the course explicitly — this has to be handled having regard to local conditions.

We have not considered it necessary to give the machine programs required in the solution of the problems: the programs are almost always trivial and when they are not, the use of library subroutines is intended. A typical problem later in Volume 2 will require, e.g., the generation of a special matrix, a call to the library for a subroutine to operate on the matrix and then a program to evaluate the error in the alleged solution provided by the machine.

Courses such as this cannot be taught properly, no matter how expert the teaching assistants are, unless the instructor has genuine practical experience in the use of computers and a minimum requirement for this is that he should have done a significant proportion of the problems himself.