The Finite Difference Method in Partial Differential Equations

A.R. Mitchell and D.F. Griffiths
Department of Mathematics,
University of Dundee, Scotland

A Wiley-Interscience Publication

JOHN WILEY & SONS
Chichester · New York · Brisbane · Toronto

British Library Cataloguing in Publication Data

Mitchell, Andrew Ronald
The finite difference method in partial differential
equations.
1. Differential equation, Partial
2. Nets (Mathematics)
1. Title II. Griffiths, D.F.
515'.353 QA374 79-40646

ISBN 0 471 27641 3

Typesetting by MHL Typesetting Ltd. Coventry
Printed by Page Brothers Limited, Norwich

Preface

The study of numerical methods for the solution of partial differential equations has enjoyed an intense period of activity over the last thirty years from both the theoretical and practical points of view. Improvements in numerical techniques, together with the rapid advances in computer technology, have meant that many of the partial differential equations from engineering and scientific applications which were previously intractable can now be routinely solved.

This book is primarily concerned with finite difference techniques and these may be thought of as having evolved in two stages. During the fifties and early sixties many general algorithms were produced and analysed for the solution of standard partial differential equations. Since then the emphasis has shifted toward the construction of methods for particular problems having special features which defy solution by more general algorithms. This approach often necessitates a greater awareness of the different physical backgrounds of the problems such as free and moving boundary problems, shock waves, singular perturbations and many others particularly in the field of fluid dynamics. The present volume attempts to deal with both aspects of finite difference development with due regard to non-linear as well as linear differential equations. Often the solution of the sparse linear algebraic equations which arise from finite difference approximations forms a major part of the problem and so substantial coverage is given to both direct and iterative methods including an introduction to recent work on preconditioned conjugate gradient algorithms.

Although finite element methods now seem to dominate the scene, especially at a research level, it is perhaps fair to say that they have not yet made the impact on hyperbolic and other time-dependent problems that they have achieved with elliptic equations. We have found it appropriate to include an introduction to finite element methods but, in the limited space available, have concentrated on their relationships to finite difference methods.

The book is aimed at final year undergraduate and first year post-graduate students in mathematics and engineering. No specialized mathematical knowledge is required beyond what is normally taught in undergraduate courses in calculus and matrix theory. Although only a rudimentary knowledge of partial differential equations is assumed, anything beyond this would seriously limit the usefulness of the book, the dangers of developing numerical methods in ignorance of the corresponding theory cannot be emphasized too strongly. Theorems and proofs of existence, uniqueness, stability and convergence are seldom given and the reader is referred to appropriate research papers and advanced texts.

The sections devoted to applications reflect the strong links with computational fluid dynamics and it is hoped that practitioners in this field will find this material useful. It is taken for granted that the reader will have access to a digital computer since we believe that a proper understanding of the many methods described, along with their limitations, will be improved greatly by practical experience.

The list of references, which we readily admit cannot do justice to the vast literature in this field, is intended for the reader who wishes to pursue the subject at greater depth. We apologize to those many authors who have made important contributions in this area and whose names have not been mentioned. Most of the material in this book has been presented in the form of lectures to Honors and M.Sc. students in the Universities of Dundee and St Andrews. An earlier version of this text, written by one of the present authors, was published under the title *Computational Methods in Partial Differential Equations,* John Wiley and Sons, 1969.

In preparing the material for this book the authors have benefited greatly from discussions with many colleagues and former students. Special thanks are due to Ian Christie, Graeme Fairweather, Roger Fletcher, Sandy Gourlay, Pat Keast, Jack Lambert, John Morris and Bill Morton. Final thanks are due to Ros Hume for her swift and accurate typing of the manuscript.

A.R. MITCHELL
D.F. DRIFFITHS

Contents

Preface xi

1 Basic Linear Algebra 1

1.1 Introduction 1
1.2 Notation 2
1.3 Definitions 2
1.4 Linear vector space 5
1.5 Useful matrix properties 7
1.6 Vector and matrix norms 10
1.7 Theorems relating matrix norm to spectral radius 13
1.8 Convergence of sequences of matrices 15

2 Parabolic Equations 17

2.1 Parabolic equations in one space dimension 17
2.2 Derivation of an exact difference formula 18
2.3 Explicit formulae 20
2.4 Implicit formulae 28
2.5 Solution of tridiagonal systems 32
2.6 Convergence 35
2.7 Stability 38
2.8 Derivative boundary conditions 47
2.9 Parabolic equations in two space dimensions 53
2.10 Explicit methods 54
2.11 Stability of explicit methods 58
2.12 Alternating direction implicit (A.D.I.) methods 59
2.13 Locally one-dimensional (L.O.D.) methods 70
2.14 Hopscotch methods 77
2.15 Mixed space derivative 81
2.16 Parabolic equations in three space dimensions 82
2.17 Explicit methods 83

2.18 Implicit methods 85
2.19 Three-level difference schemes 89
2.20 Explicit schemes 89
2.21 Implicit schemes 92
2.22 Non-linear equations 97

3 Elliptic Equations 102

3.1 Elliptic equations in two dimensions 102
3.2 Laplace's equation in a square 103
3.3 The Neumann problem 105
3.4 Mixed boundary conditions 109
3.5 Non-rectangular regions 111
3.6 Self-adjoint elliptic equations 114
3.7 Alternative methods for constructing difference formulae 116
3.8 General properties of difference formulae 121
3.9 Mixed partial derivatives 125
3.10 The biharmonic equation 129
3.11 The solution of elliptic difference equations 133
3.12 Direct factorization methods 133
3.13 Successive over-relaxation (S.O.R.) 137
3.14 A.D.I. methods 148
3.15 Conjugate gradient and related methods 154
3.16 Eigenvalue problems 161

4 Hyperbolic Equations 164

4.1 First-order hyperbolic equations 164
4.2 Explicit difference formulae 165
4.3 Implicit difference schemes 168
4.4 First-order hyperbolic systems in one space dimension 172
4.5 Systems of conservation laws 176
4.6 First-order hyperbolic systems in two space dimensions 180
4.7 Conservation laws in two space dimensions 184
4.8 Dissipation and dispersion 187
4.9 Stability of initial boundary value problems 189
4.10 Non-linear instability 191
4.11 Second-order equations in one space dimension 193
4.12 Second-order equations in two space dimensions 200

5 The Galerkin Method 210

5.1 Introduction 210
5.2 Elliptic equations 211
5.3 Two-point boundary value problems 213
5.4 The Galerkin method with different test and trial functions 221

5.5 Two space variables 223
5.6 Semi-discrete Galerkin methods 226
5.7 Discretization in time 230
5.8 Non-linear problems 232

6 Applications 238

6.1 Re-entrant corners and boundary singularities 238
6.2 Incompressible viscous flow 244
6.3 Inviscid compressible flow 249
6.4 Other flows 252
6.5 Shallow water equations 252
6.6 Free and moving boundary problems 255
6.7 Error growth in conduction-convection problems 258

References 262

Index 268

Chapter 1

Basic Linear Algebra

1.1 Introduction

In the numerical solution of partial differential equations by finite difference methods, the differential system is replaced by a matrix system where the matrix A is usually *square* with *real* elements. In the present chapter some useful properties of the matrix A are outlined, often without proof. For more detailed information concerning properties of matrices, the reader is referred to books such as Fox (1964) Wilkinson (1965) and Fadeeva (1959).

The system of linear equations requiring solution is

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \ (i = 1,2, \ldots, n), \tag{1}$$

which may be written as the matrix system

$$A\mathbf{x} = \mathbf{b}, \tag{2}$$

where A has n rows and columns and the elements $a_{ij}(i,j = 1,2, \ldots, n)$ are real numbers. The vectors $\mathbf{x}$ and $\mathbf{b}$ have n components.

The usual problem is to find $\mathbf{x}$ when A and $\mathbf{b}$ are given. A unique solution of equation (2) which may be written in the form

$$\mathbf{x} = A^{-1}\mathbf{b}$$

exists for equation (2), when A is non-singular, which is equivalent to A having a non-vanishing determinant. Since equation (2) is a matrix representation of a differential system, the matrix A is usually *sparse* (many of its elements are zero) and possesses a definite structure (determined by its non-zero elements). The method of inversion of A, particularly when the order n of the matrix is large, depends very much on the structure of A, and a variety of techniques for inverting A will be presented throughout this book. As n becomes larger, the methods of inversion must become more efficient.

A review of notation and properties for a square matrix A of order n with real elements, which is relevant to the solution of equation (1), is now given.

1.2 Notation

A	square matrix of order n.
a_{ij}	number in the real field, which is the element in the ith row and jth column of the matrix A.
A^{-1}	inverse of A.
A^T	transpose of A.
$\|A\|$	determinant of A.
$\rho(A)$	spectral radius of A.
I	unit matrix of order n.
0	null matrix.
$\mathbf{x}$	column vector with elements x_i $(i = 1,2, \ldots, n)$.
$\mathbf{x}^T$	row vector with elements x_j $(j = 1,2, \ldots, n)$.
$\bar{\mathbf{x}}$	complex conjugate of $\mathbf{x}$.
$\|\|A\|\|$	norm of A.
$\|\|\mathbf{x}\|\|$	norm of $\mathbf{x}$.
Π	permutation matrix which has entries of zeros and ones only, with one non-zero entry in each row and column.

1.3 Definitions

The matrix A is
non-singular if $|A| \neq 0$.
symmetric if $A = A^T$.
orthogonal if $A^{-1} = A^T$.
null if $a_{ij} = 0$ $(i,j = 1,2, \ldots, n)$.
diagonal if $a_{ij} = 0$ $(i \neq j)$.
diagonally dominant if $|a_{ii}| \geqslant \sum_{j \neq i} |a_{ij}|$ for all i.
tridiagonal if $a_{ij} = 0$ for $|i - j| > 1$.
block diagonal if

$$A = \begin{bmatrix} B_1 & & & & \bigcirc \\ & B_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ \bigcirc & & & & B_s \end{bmatrix}$$

where each $B_k (k = 1, 2, \ldots, s)$ is a square matrix, not necessarily of the same order.

upper triangular if $a_{ij} = 0, i > j$.
lower triangular if $a_{ij} = 0, j > i$.
irreducible if there exists no permutation transformation $\Pi A \Pi^{-1}$ which reduces A to the form

$$\begin{bmatrix} P & 0 \\ R & Q \end{bmatrix}$$

where P and Q are square submatrices of order p and q respectively ($p + q = n$) and 0 is a $p \times q$ null matrix.

The characteristic equation of A is

$$|A - \lambda I| = 0. \tag{3}$$

The eigenvalues of A are the roots $\lambda_i (i = 1, 2, \ldots, n)$ of the characteristic equation.

A right* eigenvector $\mathbf{x}^{(i)}$ for each λ_i is given by

$$A\mathbf{x}^{(i)} = \lambda_i \mathbf{x}^{(i)} \qquad (x^{(i)} \neq 0).$$

A left eigenvector $\mathbf{y}^{(i)T}$ for each λ_i is given by

$$\mathbf{y}^{(i)T} A = \lambda_i \mathbf{y}^{(i)T} \quad (\mathbf{y}^{(i)} \neq 0).$$

Two matrices A and B are similar if $B = H^{-1} AH$ for some non-singular matrix H. $H^{-1} AH$ is a similarity transformation of A.

A and B commute if $AB = BA$.

Example 1 Find the values of λ for which the set of equations

$$\begin{aligned} (1 - \lambda)x_1 + x_2 - x_3 &= 0 \\ x_1 + (2 - \lambda)x_2 + x_3 &= 0 \\ x_1 + x_2 + (3 - \lambda)\, x_3 &= 0 \end{aligned}$$

has a non-zero solution and find one such solution.

A non-zero solution exists only if

$$\begin{vmatrix} 1 - \lambda & 1 & -1 \\ 1 & 2 - \lambda & 1 \\ 1 & 1 & 3 - \lambda \end{vmatrix} = 0,$$

which leads to

$$\lambda = 2, 2 \pm \sqrt{2}.$$

* The word 'right' is often omitted and $\mathbf{x}^{(i)}$ is usually referred to an an eigenvector corresponding to λ_i.

For $\lambda = 2$, the equations become

$$\left.\begin{aligned} &\text{(a)} \quad -x_1 + x_2 - x_3 = 0, \\ &\text{(b)} \quad x_1 + x_3 = 0, \\ &\text{(c)} \quad x_1 + x_2 + x_3 = 0. \end{aligned}\right\} \tag{4}$$

Only two of the equations of (4) are independent. For example, (b) can be obtained by halving the difference of (c) and (a). If we ignore (a) it follows from (b) and (c) that

$$x_2 = 0 \text{ and } x_1/x_3 = -1.$$

Thus any eigenvector corresponding to $\lambda = 2$ is proportional to

$$\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

Example 2 If $\mathbf{x}^{(i)}$ $(i = 1,2, \ldots, n)$ *are the eigenvectors of* A *and* $\mathbf{y}^{(j)}$ $(j = 1,2, \ldots, n)$ *are the eigenvectors of* A^T, *show that*

$$\mathbf{x}^{(i)T}\mathbf{y}^{(j)} = 0 \quad (\lambda_i \neq \lambda_j),$$

where λ_i $(i = 1,2, \ldots, n)$ *are the eigenvalues of* A.

The eigenvalues of A^T are given by

$$|A^T - \lambda I| = 0,$$

since the eigenvalues of A^T are the same as those of A. An eigenvector of A^T corresponding to λ_j is $\mathbf{y}^{(j)}$, given by

$$A^T\mathbf{y}^{(j)} = \lambda_j\mathbf{y}^{(j)}, \tag{5}$$

or after transposing both sides,

$$\mathbf{y}^{(j)T}A = \lambda_j\mathbf{y}^{(j)T}.$$

($\mathbf{y}^{(j)T}$ is a left eigenvector of A.)

Also

$$A\mathbf{x}^{(i)} = \lambda_i\mathbf{x}^{(i)},$$

which, on transposing both sides, gives

$$\mathbf{x}^{(i)T}A^T = \lambda_i\mathbf{x}^{(i)T}. \tag{6}$$

Postmultiplying equation (6) by $\mathbf{y}^{(j)}$ and premultiplying equation (5) by $\mathbf{x}^{(i)T}$, and subtracting, the result

$$0 = (\lambda_i - \lambda_j)\mathbf{x}^{(i)T}\mathbf{y}^{(j)}$$

is obtained. This leads to the desired result if $\lambda_i \neq \lambda_j$.

Example 3 Show that the eigenvalues of a matrix are preserved under a similarity transformation.

If

$$Ax = \lambda x, \qquad x \neq 0$$

then

$$H^{-1} Ax = \lambda H^{-1} x \quad (|H| \neq 0),$$

leading to

$$H^{-1} AHH^{-1} x = \lambda H^{-1} x,$$

and so

$$(H^{-1} AH)H^{-1} x = \lambda H^{-1} x.$$

Thus the eigenvalues are preserved and the eigenvectors are multiplied by H^{-1}.

Example 4 If the eigenvalues of the matrix A are distinct, show that there is a similarity transformation which reduces A to diagonal form and has its columns equal to the eigenvectors of A.

A has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and eigenvectors $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$, A^T has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and eigenvectors $y^{(1)}, y^{(2)}, \ldots, y^{(n)}$, where

$$y^{(j)T} x^{(i)} = 0 \quad (i \neq j),$$

and

$$y^{(i)T} x^{(i)} = 1 \quad (i = 1, 2, \ldots, n),$$

the eigenvectors having been normalized. These relations imply that the matrix Y^T which has $y^{(j)T}$ as its jth row is the inverse of the matrix X, which has $x^{(i)}$ as its ith column. Now

$$AX = X \operatorname{diag}(\lambda_i),$$

and since Y^T is the inverse of X, it follows that

$$X^{-1} AX = Y^T AX = \operatorname{diag}(\lambda_i),$$

leading to the desired result.

1.4 Linear vector space

An important role will be played by the n-dimensional vector space R_n. A point x in this space is arrayed in the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix},$$

where $x_1, x_2, \ldots, x_n$ are the n components of the vector, which may be complex, but which we shall assume to be real. The number n is said to be the *dimension* of the space.

Vectors $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$ are said to be *linearly dependent* if non-zero constants $c_1, c_2, \ldots, c_n$ exist such that

$$c_1\mathbf{x}^{(1)} + c_2\mathbf{x}^{(2)} + \ldots + c_n\mathbf{x}^{(n)} = 0.$$

If this equation holds only for $c_1 = c_2 = \ldots = c_n = 0$, however, the vectors $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$ are said to be *linearly independent.* A system of linearly independent vectors is said to constitute a *basis* for a space, if any vector of the space is a linear combination of the vectors of the system. The number of vectors forming a basis is equivalent to the dimension of the space. The n linearly independent vectors form a complete system and are said to span the whole n space.

The *inner* (*or scalar*) *product* of two members $\mathbf{x}$ and $\mathbf{y}$ of the vector space is defined by

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} x_i y_i.$$

The *length* of a vector $\mathbf{x}$ is given by

$$(\mathbf{x}, \mathbf{x})^{1/2} = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}$$

The non-zero vectors $\mathbf{x}$ and $\mathbf{y}$ are said to be *orthogonal* if $(\mathbf{x}, \mathbf{y}) = 0$. A system of vectors is orthogonal if any two vectors of the system are orthogonal to one another.

Theorem The vectors forming an orthogonal system are linearly independent.

Proof: Let $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$ form an orthogonal system and suppose that

$$c_1\mathbf{x}^{(1)} + c_2\mathbf{x}^{(2)} + \ldots + c_n\mathbf{x}^{(n)} = 0.$$

Take the scalar product with $\mathbf{x}^{(i)}$, and so

$$c_i(\mathbf{x}^{(i)}, \mathbf{x}^{(i)}) = 0$$

for any $i = 1, 2, \ldots, n$. Since $(\mathbf{x}^{(i)}, \mathbf{x}^{(i)}) \neq 0$, it follows that

$$c_i = 0 \quad (i = 1, 2, \ldots, n).$$

Thus the vectors $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$ are linearly independent.

1.5 Useful matrix properties

As far as possible these properties are grouped, although there is no particular merit in the order of presentation chosen. All matrices are square of order n. We shall be concerned almost entirely with real matrices, but some of the results in this section apply when A is complex. This is particularly true of the Jordan canonical form where one usually requires to work with complex numbers.

Jordan canonical form

A Jordan submatrix of A is a matrix of the form

$$\begin{bmatrix} \lambda_i & & & & \bigcirc \\ 1 & \lambda_i & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ \bigcirc & & & 1 & \lambda_i \end{bmatrix}$$

where λ_i is an eigenvalue of A. The Jordan canonical form of A is a block diagonal matrix composed of Jordan submatrices. It is unique up to permutations of the blocks. Any matrix A can be reduced to Jordan canonical form by a similarity transformation

$$J = H^{-1}AH.$$

The diagonal elements of J are the eigenvalues of A.

If A has n distinct eigenvalues, its Jordan canonical form is diagonal and its n associated eigenvectors are linearly independent. They form a complete system of eigenvectors and span the whole n-dimensional space. If A does not have n distinct eigenvalues, it may or may not possess n independent eigenvectors.

If any two matrices A and B commute and have diagonal canonical forms, then they have a complete set of simultaneous eigenvectors.

Symmetric matrix

A symmetric matrix has

(i) a diagonal Jordan canonical form;
(ii) n real eigenvalues; and
(iii) n mutually orthogonal eigenvectors.

If A and B are symmetric, and $AB = BA$, then AB is symmetric.

Positive definite matrix

If A is real and $\mathbf{x}$ is complex, then A is *positive definite* if

$$(\mathbf{x}, A\mathbf{x}) > 0 \text{ for all } \mathbf{x} \neq 0.$$

This, of course, implies that $(\mathbf{x}, A\mathbf{x})$ is real. If A is positive definite, then it is symmetric. (Note that the inner product $(\mathbf{x}, \mathbf{y})$ of two complex vectors is $\sum_{i=1}^{n} x_i \bar{y}_i$, where $\bar{y}_i$ is the complex conjugate of y_i)

If A is real and $\mathbf{x}$ is real, then A is *positive real* if

$$(\mathbf{x}, A\mathbf{x}) > 0 \text{ for all } \mathbf{x} \neq 0.$$

This time A is not necessarily symmetric.

A matrix A is *positive semi-definite* if

$$(\mathbf{x}, A\mathbf{x}) \geqslant 0 \text{ for all } \mathbf{x},$$

with equality for at least one $\mathbf{x} \neq 0$.

A *Stieltjes matrix* is a real positive definite matrix with all its off-diagonal elements non-positive. If the properties of irreducibility and diagonal dominance are added, the matrix is often referred to as an S-matrix. An S-matrix has the following properties:

(i) $a_{ij} = a_{ji}$;
(ii) $a_{ij} \leqslant 0$ for $i \neq j$;
(iii) $|a_{ii}| \geqslant \sum_{j \neq i} |a_{ij}|$, with strict inequality for at least one i;
(iv) $S \equiv [a_{ij}]$ is irreducible;
(v) S is positive definite; and
(vi) the elements of S^{-1} are positive.

Such matrices occur repeatedly in the finite difference solution of partial differential equations.

Example 5 Show that, if $(\mathbf{x}, A\mathbf{x}) > 0$ *for all complex* $\mathbf{x}$, *then* A *is symmetric.*

Let $\mathbf{x} = \mathbf{a} + i\mathbf{b}$ where $\mathbf{a}$ and $\mathbf{b}$ are real. Now the inner product $(\mathbf{x}, \mathbf{y})$ of two complex vectors is

$$\sum_{i=1}^{n} x_i \bar{y}_i,$$

where $\bar{y}_i$ is the complex conjugate of y_i. Consequently,

$$\begin{aligned}(\mathbf{x}, A\mathbf{x}) &= (\mathbf{a} + i\mathbf{b}, A(\mathbf{a} + i\mathbf{b})) \\ &= (\mathbf{a}, A\mathbf{a}) + i(\mathbf{b}, A\mathbf{a}) - i(\mathbf{a}, A\mathbf{b}) + (\mathbf{b}, A\mathbf{b}) \\ &= [(\mathbf{a}, A\mathbf{a}) + (\mathbf{b}, A\mathbf{b})] - i[(\mathbf{a}, A\mathbf{b}) - (\mathbf{b}, A\mathbf{a})] > 0.\end{aligned}$$

This is only possible if

$$(\mathbf{a}, A\mathbf{b}) - (\mathbf{b}, A\mathbf{a}) = (\mathbf{a}, A\mathbf{b}) - (\mathbf{a}, A^T\mathbf{b}) = (\mathbf{a}, (A - A^T)\,\mathbf{b}) = 0,$$

and so

$$A = A^T,$$

leading to the desired result.

Example 6 Show that if A is symmetric and positive real, then its eigenvalues are all positive.

If A is symmetric, its real eigenvalues imply real eigenvectors. Consequently,

$$(\mathbf{x}, A\mathbf{x}) = (\mathbf{x}, \lambda\mathbf{x}) = \lambda(\mathbf{x}, \mathbf{x})$$

for any eigenvector $\mathbf{x} \neq 0$. But

$$(\mathbf{x}, A\mathbf{x}) > 0,$$

since A is positive real, and so

$$\lambda = \frac{(\mathbf{x}, A\mathbf{x})}{(\mathbf{x}, \mathbf{x})} > 0.$$

Example 7 Show that A^TA has non-negative eigenvalues.

Let

$$B = A^TA.$$

B is symmetric because

$$B^T = (A^TA)^T = A^TA = B,$$

and so B has real eigenvalues and real eigenvectors. For any real non-zero $\mathbf{x}$,

$$(\mathbf{x}, B\mathbf{x}) = (\mathbf{x}, A^TA\mathbf{x}) = (A\mathbf{x}, A\mathbf{x}) \geqslant 0,$$

and so B is positive semi-real and, by previous example, has non-negative eigenvalues.

Eigenvalues of a matrix

The eigenvalues of A lie within the union of the n discs

$$|z - a_{ii}| \leqslant \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \quad (i = 1, 2, \ldots, n)$$

in the complex z plane. Since A^T has the same eigenvalues as A, this may be replaced by

$$|z - a_{jj}| \leqslant \sum_{\substack{i=1 \\ i \neq j}}^{n} |a_{ij}| \qquad (j = 1, 2, \ldots, n).$$

This is *Gerschgorin's theorem.*

The *spectral radius* of a matrix A is denoted by $\rho(A)$ and is given by

$$\rho(A) = \max_i |\lambda_i|,$$

where $\lambda_i (i = 1, 2, \ldots, n)$ are the eigenvalues of A. $\rho(A)$ is the radius of the smallest circular disc in the complex plane, with the centre as the origin, which contains all the eigenvalues of A. From Gerschgorin's theorem,

$$\rho(A) \leqslant \min \left(\max_i \sum_j |a_{ij}|; \max_j \sum_i |a_{ij}| \right).$$

If A is the tridiagonal matrix

$$\begin{bmatrix} a & b & & & & \bigcirc \\ c & a & b & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & b \\ \bigcirc & & & & c & a \end{bmatrix}$$

where a, b and c are real and $bc > 0$, the eigenvalues of A are given in closed form by

$$\lambda_s = a + 2\surd(bc) \cos \frac{s\pi}{n+1} \quad (s = 1, 2, \ldots, n).$$

Exercise

1. Find the eigenvectors of the matrix

$$\begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix},$$

and show that they are mutually orthogonal.

1.6 Vector and matrix norms

The modulus of a complex number gives an assessment of its overall size. It will