J. SIAM NUMER. ANAL.
Vol. 3, No. 2, 1966
Printed in U.S.A.

ON A DISCRETE MAXIMUM PRINCIPLE*

RICHARD S. VARGA†

1. Introduction. Recently, some higher-order difference methods [1], [2], [4] have been advocated for numerically solving second-order elliptic boundary value problems such as Laplace's equation, and the question naturally arises as to whether the associated discrete problems satisfy, like the continuous problem, a *maximum principle*, i.e., the maximum component in modulus of the solution vector is bounded above by the maximum component in modulus of the boundary data. For the simplest difference approximations, this is easy to answer affirmatively. Because each unknown in this case is the average of some of its neighboring unknowns (cf. (4)), the well-known proof by contradiction for the continuous case directly carries over to the discrete case. For higher-order difference approximations, the answer is not immediately obvious since this averaging property is lost (cf. (7)).

In §2 of this paper we give necessary and sufficient conditions that a matrix satisfy a discrete maximum principle with respect to a given subspace, and in §3 we apply these results to the difference methods of [1], [2], and show that the associated discrete problems do indeed satisfy a discrete maximum principle.

2. Let $V_n(C)$ denote the n-dimensional vector space of column vectors $\mathbf{x} = [x_1, \cdots, x_n]^T$ having complex components. We can express $\mathbf{x}$ as $\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$, where

$$\{\mathbf{e}_i = [\delta_{1,i}, \delta_{2,i}, \cdots, \delta_{n,i}]^T\}_{i=1}^{n}$$

is an orthonormal basis for $V_n(C)$. As usual, $\| \mathbf{x} \|_\infty$ is defined as $\max_{1 \leq i \leq n} | x_i |$ for any $\mathbf{x} \in V_n(C)$, and if $B = (b_{i,j})$ is any $n \times n$ (complex) matrix relative to this basis, the associated operator norm of B is given by

$$\| B \|_\infty = \sup_{\| \mathbf{x} \|_\infty = 1} \| B\mathbf{x} \|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} | b_{i,j} |. \tag{1}$$

Let S denote any r-dimensional linear subspace of $V_n(C)$ spanned by r of the vectors $\mathbf{e}_j$, $1 \leq r \leq n$. Associated with the subspace S is the $n \times n$ *projection matrix* $P_s = \text{diag}\,(d_1, d_2, \cdots, d_n)$, where $d_i = 1$ if $\mathbf{e}_i \in S$, and zero otherwise. Thus, $P_s\mathbf{x} \in S$ and $\| P_s\mathbf{x} \|_\infty \leq \| \mathbf{x} \|_\infty$ for any $\mathbf{x} \in V_n(C)$, so that $\| \mathbf{P}_s \| \leq 1$.

DEFINITION 1. An $n \times n$ matrix A satisfies a *discrete maximum principle*

* Received by the editors November 15, 1965.
Dedicated to Professor J. L. Walsh on the occasion of his seventieth birthday.
† Computing Center, Case Institute of Technology, Cleveland, Ohio.

with respect to the subspace S (written $A \in \mathfrak{N}_s$) if and only if, given any $\mathbf{g} \in V_n(C)$, every solution $\mathbf{u}$ of $A\mathbf{u} = P_s\mathbf{g}$ satisfies $\| \mathbf{u} \|_\infty \leqq \| P_s\mathbf{g} \|_\infty$.

We remark that if $A \in \mathfrak{N}_s$, then choosing $\mathbf{g} = \mathbf{0}$ in $V_n(C)$ implies that the solution $\mathbf{u}$ of $A\mathbf{u} = \mathbf{0}$ is necessarily $\mathbf{u} = \mathbf{0}$. Thus, A is nonsingular.

LEMMA. *$A \in \mathfrak{N}_s$ if and only if A is nonsingular and $\| A^{-1}P_s \|_\infty \leqq 1$.*

Proof. If A is nonsingular, then the solution of $A\mathbf{u} = P_s\mathbf{g}$ is $\mathbf{u} = A^{-1}P_s\mathbf{g} = A^{-1}P_sP_s\mathbf{g}$, since $P_s^2 = P_s$. Taking norms, $\| \mathbf{u} \|_\infty \leqq \| A^{-1}P_s \|_\infty \cdot \| P_s\mathbf{g} \|_\infty$. If $\| A^{-1}P_s \|_\infty \leqq 1$, then $\| \mathbf{u} \|_\infty \leqq \| P_s\mathbf{g} \|_\infty$ for any $\mathbf{g} \in V_n(C)$, so that $A \in \mathfrak{N}_s$. Conversely, if $A \in \mathfrak{N}_s$, then A is nonsingular, and the solution $\mathbf{u}$ of $A\mathbf{u} = P_s\mathbf{g}$ satisfies $\| \mathbf{u} \|_\infty = \| A^{-1}P_s\mathbf{g} \|_\infty \leqq \| P_s\mathbf{g} \|_\infty \leqq \| \mathbf{g} \|_\infty$ for any $\mathbf{g} \in V_n(C)$, and hence $\| A^{-1}P_s \|_\infty \leqq 1$, completing the proof.

DEFINITION 2. An $n \times n$ matrix A is *normalized with respect to the subspace S* (written $A \in \mathfrak{N}_s$) if and only if $A\xi = P_s\xi$, where $\xi = \sum_{i=1}^{n} \mathbf{e}_i$.

In other words, if $A = (a_{i,j})$ is the matrix relative to the basis $\{\mathbf{e}_i\}_{i=1}^{n}$, then the row sum $\sum_{j=1}^{n} a_{i,j}$ is, respectively, unity or zero if $\mathbf{e}_i \in S$ or $\mathbf{e}_i \notin S$. Writing $B \geqq 0$ if B is an $n \times n$ matrix with nonnegative real entries, we then prove the following.

THEOREM. *Let $A \in \mathfrak{N}_s$. Then $A \in \mathfrak{N}_s$ if and only if A is nonsingular and $A^{-1}P_s \geqq 0$.*

Proof. If A is nonsingular, then as $A \in \mathfrak{N}_s$, $\xi = A^{-1}P_s\xi$, so that the row sums of $A^{-1}P_s$ are all unity. If $A^{-1}P_s \geqq 0$ in addition, then the row sums of the moduli of the entries of $A^{-1}P_s$ are also unity, so that $\| A^{-1}P_s \|_\infty = 1$ from (1). From the Lemma, $A \in \mathfrak{N}_s$. Conversely, if $A \in \mathfrak{N}_s$, then A is nonsingular and $\| A^{-1}P_s \|_\infty \leqq 1$. But as $A \in \mathfrak{N}_s$, then $\xi = A^{-1}P_s\xi$ shows that $\| A^{-1}P_s \|_\infty \geqq 1$. Combining, $\| A^{-1}P_s \|_\infty = 1$. Hence, the row sums of $A^{-1}P_s$ are all unity from $\xi = A^{-1}P_s\xi$, and from (1), the maximum of the row sums of the moduli of the entries of $A^{-1}P_s$ is unity since $\| A^{-1}P_s \|_\infty = 1$. It necessarily follows that $A^{-1}P_s \geqq 0$, which completes the proof.

Since the projection P_s is itself a nonnegative matrix relative to the basis $\{\mathbf{e}_i\}_{i=1}^{n}$, we have as an immediate consequence of this theorem:

COROLLARY 1. *If $A \in \mathfrak{N}_s$, and A is nonsingular with $A^{-1} \geqq 0$, then $A \in \mathfrak{N}_s$.*

With the notation $| A | = (| a_{i,j} |)$ if $A = (a_{i,j})$, we prove another consequence of this theorem.

COROLLARY 2. *If $B \in \mathfrak{N}_s \cap \mathfrak{N}_s$ and A is an $n \times n$ nonsingular matrix with $| A^{-1}P_s | \leqq B^{-1}P_s$, then $A \in \mathfrak{N}_s$.*

Proof. From the proof of the theorem, $\| B^{-1}P_s \|_\infty = 1$, and from (1), it follows that $\| A^{-1}P_s \|_\infty = \| | A^{-1}P_s | \|_\infty \leqq \| B^{-1}P_s \|_\infty = 1$, so that $A \in \mathfrak{N}_s$ from Lemma 1.

An $n \times n$ matrix A for which A is nonsingular and $A^{-1} \geqq 0$ is called a *monotone* matrix [3]. From Corollary 1 above, it is clear that if $A \in \mathfrak{N}_s$, then

A monotone is sufficient for $A \in \mathfrak{M}_s$. This condition, however, is not necessary. As a simple example, let $n = 3$, $r = 1$, let S be spanned by $\mathbf{e}_1$, and choose A as given below:

$$A = \begin{bmatrix} -1 & 2 & 0 \\ 2 & -3 & 1 \\ 0 & 1 & -1 \end{bmatrix}, \qquad A^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \frac{1}{2} & \frac{1}{2} \\ 1 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}, \qquad P_s = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Clearly, $A \in \mathfrak{N}_s$ and $A^{-1}P_s \geqq 0$ so that $A \in \mathfrak{N}_s \cap \mathfrak{M}_s$, but A is not monotone.

3. We now apply our previous results to the numerical solution of elliptic boundary value problems. First, we consider the numerical solution of

$$-u_{xx}(x) = 0, \qquad 0 < x < 1, \tag{2}$$

subject to the boundary conditions

$$u(0) = \alpha, \qquad u(1) = \beta. \tag{3}$$

Here, α and β are given. With $h \equiv 1/(N+1)$, $x_i = ih$, $0 \leqq i \leqq N+1$, we can approximate the solution of (2)–(3) by the following system of linear difference equations:

$$\begin{aligned} &(Aw)_1 = w_1 = \alpha, \\ &(Aw)_i = \frac{-w_{i-1} + 2w_i - w_{i+1}}{h^2} = 0 \qquad \text{if } 2 \leqq i \leqq N+1, \\ &(Aw)_{N+2} = w_{N+2} = \beta. \end{aligned} \tag{4}$$

Here, A is an $(N+2) \times (N+2)$ real tridiagonal matrix which is easily verified to be an M-matrix [5, p. 84], so that A is nonsingular and $A^{-1} \geqq 0$. If S is the linear subspace of $V_{N+2}(C)$ spanned by the vectors $\mathbf{e}_1$ and $\mathbf{e}_{N+2}$, then it is obvious from (4) that $A \in \mathfrak{N}_s$. Thus, applying Corollary 1 gives us that $A \in \mathfrak{M}_s$.

As an application of Corollary 2, consider the numerical solution of

$$-u_{xx}(x) + \sigma(x)u(x) = 0, \qquad 0 < x < 1, \tag{5}$$

subject to the boundary conditions of (3), where $\sigma(x)$ is continuous and nonnegative in $0 \leqq x \leqq 1$. We can approximate the solution of (5)–(3) by the following system of linear difference equations:

$$\begin{aligned} &(Bw)_1 = w_1 = \alpha, \\ &(Bw)_i = \frac{-w_{i-1} + (2 + \sigma_{i-1}h^2)w_i - w_{i+1}}{h^2} = 0 \qquad \text{if } 2 \leqq i \leqq N+1, \\ &(Bw)_{N+2} = w_{N+2} = \beta, \end{aligned} \tag{6}$$

where $\sigma_i = \sigma(x_i)$. Since $\sigma_i \geqq 0$, then B is an $(N+2) \times (N+2)$ tridiagonal M-matrix, and as such, B is nonsingular with $B^{-1} \geqq 0$. With S defined again as the linear subspace of $V_{N+2}(C)$ spanned by $\mathbf{e}_1$ and $\mathbf{e}_{N+2}$, it is not now in general true that $B \in \mathfrak{N}_s$. On the other hand, as $B \geqq A$, and $B^{-1} \geqq 0$ and $A^{-1} \geqq 0$, it then follows [5, p. 87] that $A^{-1} \geqq B^{-1}$, where A is defined in (4). Thus, $|B^{-1}P_s| \leqq A^{-1}P_s$, and from Corollary 2, we deduce that $B \in \mathfrak{M}_s$.

The approximations used in (4) are $O(h^2)$ in the sense that if $u(x)$, the solution of (2), is an element of $C^4[0, 1]$, then $\max_{1 \leqq i \leqq N+2} |u(x_i) - w_{i+1}| = O(h^2)$, as $h \to 0$. The same is true of the approximations of (6). To derive $O(h^4)$ difference approximations of (2)–(3) under the assumption that $u(x) \in C^6[0, 1]$, we consider, as in [2], the system of linear difference equations:

$$
\begin{aligned}
&(Dw)_1 = w_1 = \alpha, \\
&(Dw)_2 = \frac{-w_1 + 2w_2 - w_3}{h^2} = f_1, \\
&(Dw)_i = \frac{30w_i - 16(w_{i+1} + w_{i-1}) + (w_{i+2} + w_{i-2})}{12h^2} = 0 \qquad \text{if } 3 \leqq i \leqq N, \\
&(Dw)_{N+1} = \frac{-w_N + 2w_{N+1} - w_{N+2}}{h^2} = 0, \\
&(Dw)_{N+2} = w_{N+2} = \beta.
\end{aligned} \tag{7}
$$

It is known [2] that the $(N+2) \times (N+2)$ matrix D is monotone, i.e., D is nonsingular and $D^{-1} \geqq 0$. Letting S again denote the linear subspace of $V_{N+2}(C)$ spanned by $\mathbf{e}_1$ and $\mathbf{e}_{N+2}$, then it follows from (7) that $D \in \mathfrak{N}_s$. Hence, by Corollary 1, $D \in \mathfrak{M}_s$. In other words, *the higher-order difference approximations of* [1], [2] *do satisfy a discrete maximum principle.*

We have considered in detail these ideas specifically for one-dimensional problems because they extend easily to higher dimension. To show this, consider now an $O(h^4)$ difference approximation to

$$u_{xx}(x, y) + u_{yy}(x, y) = 0, \qquad (x, y) \in R, \tag{8}$$

subject to

$$u(x, y) = g(x, y), \qquad (x, y) \in \partial R, \tag{9}$$

is a bounded region R with boundary ∂R. The mesh points of $R + \partial R$ are, as in (7), separated into three sets. The first of these sets consists of all boundary mesh points for which u is specified. At such points, we simply write down the elementary difference equation $w_l = g(x_i, y_j)$. In one dimension, this set corresponds to the end mesh points, and the difference

equations from (7) are $w_1 = \alpha$ and $w_{N+2} = \beta$. The second set of mesh points are those interior mesh points adjacent to the boundary. At such mesh points, the simplest difference equation approximating (8), but involving at most five unknowns, is used. The mesh lengths at such points are not in general equal, but the sum of the coefficients in these difference equations is nevertheless zero (cf. [5, p. 186]), basically because these difference equations are obtained from Taylor series expansions of the solution u at these mesh points. In one dimension, this set corresponds to the mesh points $i = 2$ and $i = N + 1$. The final set of mesh points is the remaining set of interior mesh points where a higher-order difference approximation to (8) involving nine mesh points is used. Again, for this set, the sum of the coefficients in the difference equation for each point in the set is zero. Thus, if we let S be the linear subspace spanned by the vectors corresponding to boundary mesh points, then the discussion above shows that the matrix arising from these finite difference approximations is an element of $\mathfrak{N}_s$. Moreover, this matrix is monotone [1], [2], [4], so that we finally conclude from Corollary 1 that this matrix is an element of $\mathfrak{M}_s$.

REFERENCES

[1] J. H. Bramble and B. E. Hubbard, *On the formulation of finite difference analogues of the Dirichlet problem for Poisson's equation*, Numer. Math., 4 (1962), pp. 313–327.

[2] ———, *On a finite difference analogue of an elliptic boundary value problem which is neither diagonally dominant nor of non-negative type*, J. Math. and Phys., 43 (1964), pp. 117–132.

[3] L. Collatz, *Aufgaben monotoner Art*, Arch. Math., 3 (1952), pp. 366–376.

[4] H. S. Price, *Monotone and oscillation matrices applied to finite difference approximations*, Ph.D. Thesis, Case Institute of Technology, Cleveland, 1965.

[5] R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, New Jersey, 1962.