BLOCK DIAGONALLY DOMINANT MATRICES AND GENERALIZATIONS OF THE GERSCHGORIN CIRCLE THEOREM

DAVID G. FEINGOLD AND RICHARD S. VARGA

BLOCK DIAGONALLY DOMINANT MATRICES AND GENERALIZATIONS OF THE GERSCHGORIN CIRCLE THEOREM

DAVID G. FEINGOLD AND RICHARD S. VARGA

alizations of the nonsingular. D is a nonsingular diagonal matrix, then Householder [7] shows $\|\,D^{-1}C\,\|<1$ in some matrix norm is sufficient to conclude that establish the nonsingularity of A. exclusion regions for the eigenvalues of an arbitrary square matrix A Introduction. such exclusion regions arise naturally Hence, the set of all complex numbers z for which well known theorem of Gerschgorin on inclusion or The main purpose of this paper is to give gener-For example, if A = D + C where from results which

$$||(zI-D)^{-1}C|| < 1$$

nonsingularity of A through comparisons with M-matrices. Our approach, obtains exclusion regions for the eigenvalues of A by establishing the evidently contains no eigenvalues of A. In a like manner, Gerschgorin circles in providing bounds for the eigenvalues of A. exclusion regions can give significant improvements over dominant matrix. matrix A by the generalization of though not fundamentally different, establishes the nonsingularity of the But one of our major results (§ 3) is that these new the simple concept of a Fiedler the usual diagonally

with complex entries, which is partitioned in the following manner: Block diagonally dominant matrices. Let A be any $n \times n$ matrix

(2.1)
$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,N} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,N} \\ \vdots & & & \vdots \\ A_{N,1} & A_{N,2} & \cdots & A_{N,N} \end{bmatrix}$$

For reasons to appear in § 3, the particular choice N=1 of where the diagonal submatrices $A_{i,i}$ are square of order n_i , $1 \le i \le N$.

$$(2.1') A = [A_{1,1}]$$

of the n_i -dimensional vector subspace Ω_i into itself, we associate will be useful. Viewing the square matrix $A_{i,i}$ as a linear transformation this subspace the vector norm $||x||_{a_i}$, i.e., if x and y are elements of

Received April 11, 1962.

For the definition of an M-matrix, see § 4 or [8].

 Q_i , then

(2.2)
$$\begin{cases} ||\mathbf{x}||_{\varrho_{i}} > 0 \text{ unless } \mathbf{x} = \mathbf{0}; \\ ||\alpha \mathbf{x}||_{\varrho_{i}} = |\alpha| ||\mathbf{x}||_{\varrho_{i}} \text{ for any scalar } \alpha; \\ (||\mathbf{x} + \mathbf{y}||_{\varrho_{i}} \le ||\mathbf{x}||_{\varrho_{i}} + ||\mathbf{y}||_{\varrho_{i}}, \quad 1 \le i \le N. \end{cases}$$

norm $||A_{i,j}||$ is defined as usual by different subspaces Ω_i . Now, similarly considering the rectangular matrix $A_{i,j}$ for any $1 \le i, j \le N$ as a linear transformation from Ω_j to Ω_i , the The point here is that we can associate different vector norms with

(2.3)
$$||A_{i,j}|| \equiv \sup_{\mathbf{x} \in a_j, \mathbf{x} \neq o} \frac{||A_{i,j}\mathbf{x}||_{a_j}}{||\mathbf{x}||_{a_j}}.$$

are 1×1 matrices and $||x||_{a_i} \equiv |x|$, then the norms $||A_{i,j}||$ are just the moduli of the single entries of these matrices. As no confusion arises, we shall drop the subscripts on the different vector norms.2 Note that if the partitioning in (2.1) is such that all the matrices $A_{i,j}$

If the diagonal submatrices $A_{j,j}$ are nonsingular, and if Definition 1. Let the $n \times n$ matrix A be partitioned as in (2.1).

$$(2.4) \qquad \qquad (||A_{j,j}^{-1}||)^{-1} \geqq \sum\limits_{\substack{k=1 \\ k \neq j}}^{N} ||A_{j,k}|| \quad \text{for all } 1 \leqq j \leqq N \,,$$

strictly diagonally dominant, relative to the partitioning of (2.1). then A is block diagonally dominant, relative to the partitioning (2.1). If strict inequality in (2.4) is valid for all $1 \le j \le N$, then A is block

side of (2.4) can also be characterized form (2.3) by It is useful to point out that the quantity appearing on the lefthand

(2.5)
$$(||A_{j,j}^{-1}||)^{-1} = \inf_{\mathbf{x} \in \mathcal{Q}_{j}, \mathbf{x} \neq O} \left(\frac{||A_{j,j}\mathbf{x}||}{||\mathbf{x}||} \right),$$

whenever $A_{i,j}$ is nonsingular. With (2.5), we can then define $(||A_{i,j}^{-1}||)^{-1}$ by continuity to be zero whenever $A_{i,j}$ is singular.

||x|| = |x|, then (2.4) can be written as In the special case that all the matrices $A_{i,j}$ are 1×1 matrices and

$$(2.4') \qquad |A_{j,j}| \geq \sum\limits_{\substack{k=1\\k\neq j\\k\neq j}}^N |A_{j,k}| \quad \text{for all } 1 \leq j \leq N\,,$$

which is the usual definition of diagonal dominance

consider the case n=4, N=2 of As an example of a matrix which is block strictly diagonally dominant,

² Later, we shall use the notation $||x||_p$ to denote the l_p -norm $||x||_p \equiv (\sum_i |x_i|^p)^{1/p}$.

$$A = \begin{vmatrix} 0 & 1 & 0 & 0 \\ 2/3 & 0 & 1/3 & 0 \\ 0 & 1/3 & 0 & 2/3 \\ 0 & 0 & 1 & 0 \end{vmatrix}$$

where we choose the vector norms $||x||_{\infty} \equiv \max_{j} |x_{j}|$. In this case,

$$(\|A_{1,1}^{-1}\|)^{-1}=(\|A_{2,2}^{-1}\|)^{-1}=rac{2}{3}$$
, and $\|A_{1,2}\|=\|A_{2,1}\|=rac{1}{3}$.

Obviously, A is not diagonally dominant in the sense of (2.4').

Definition 2. The $n \times n$ partitioned matrix A of (2.1) is block irreducible if the $N \times N$ matrix $B = (b_{i,j} \equiv ||A_{i,j}||)$, $1 \leq i,j \leq N$, is irreducible, i.e., the directed graph of B is strongly connected.

法

diagonally dominant, or if A is block irreducible and block diagonally is nonsingular. dominant with inequality holding in (2.4) for at least one j, then A THEOREM 1. If the partitioned matrix A of (2.1) is block strictly

nonzero vector diagonally dominant. easy, so we consider for simplicity only the case when A is block strictly block diagonally dominant with strict inequality Proof.The extension to the case where A is block irreducible and W with Suppose that A is singular, i.e., there exists a for at least one j is

$$(2.7) A \begin{vmatrix} W_1 \\ \vdots \\ W_{\pi} \end{vmatrix} = 0$$

of (2.1). here, we have partitioned W conformally with respect to the partitioning But this is equivalent to

(2.8)
$$\sum_{\substack{j=1 \ j \neq i}}^{N} A_{i,j} W_j = -A_{i,i} W_i$$
 , $1 \le i \le N$.

where $1 \le r \le N$. Since W is a nonzero vector, normalize W so that $||W_j|| \le 1$ for all $1 \le j \le N$, and assume that equality is valid for some r, i.e., $||W_r|| = 1$ Thus, from (2.3)

see for example [6]. where C and E are square nonvoid submatrices. ³ Equivalently, there exists no $N \times N$ permutation matrix P such that $PBP^T = \begin{bmatrix} 0 & D \\ 0 & E \end{bmatrix}$ For strongly connected directed graphs

$$(2.8') \quad ||A_{r,r}W_r|| = ||\sum_{\substack{j=1\\j\neq r}}^N A_{r,j}W_j|| \leq \sum_{\substack{j=1\\j\neq r}}^M ||A_{r,j}|| \cdot |||W_j|| \leq \sum_{\substack{j=1\\j\neq r}}^N ||A_{r,j}|| \cdot ||$$

But as $A_{r,r}$ is nonsingular by hypothesis, then putting $A_{r,r}W_r =$

$$||A_{r,r}W_r|| = \frac{||A_{r,r}W_r||}{||W_r||} = \frac{||Z_r||}{||A_{r,r}^{-1}Z_r||} \ge (||A_{r,r}^{-1}||)^{-1},$$

assumption (2.4) that A is block strictly using (2.3). completes the proof for the block strictly diagonally dominant case. This combined with (2.8') gives a contradiction to the diagonally dominant, which

[10, Theorem 3, p. 185], and Fiedler [4]. to this result in the case that all the matrices $A_{i,j}$ of (2.1) are 1×1 matrices and $||x|| \equiv |x|$. It should be pointed out that the result of well known Hadamard theorem on determinants, since Theorem 1 reduces Theorem 1 itself is a special case of a more general result by Ostrowski Actually, we can regard Theorem 1 as the block analogue of the

identity matrix, suppose that a block analogue of the Gerschgorin Circle Theorem. identity matrix which is partitioned as in (2.1), and I_j is the $n_j \times n_j$ As stated in the introduction, the above theorem leads naturally to If I is the $n \times n$

$$(2.9) \qquad (\|(A_{j,j}-\lambda I_j)^{-1}\|)^{-1} > \sum\limits_{\substack{k=1\\k\neq j}}^N \|A_{j,k}\| \quad \text{for all } 1 \leq j \leq N \,.$$

dominant, which gives us λ is an eigenvalue of A, then $A-\lambda I$ cannot be block strictly diagonally Thus, we have from Theorem 1 that $A - \lambda I$ is nonsingular.

 λ of A satisfies THEOREM 2. For the partitioned matrix A of (2.1), each eigenvalue

$$(||(A_{j,j} - \lambda I_j)^{-1}||)^{-1} \leq \sum_{\substack{k=1\\k \neq j}}^{N} ||A_{j,k}||$$

for at least one j, $1 \le j \le N$.

reduces to the well known Gerschgorin Circle Theorem. the diagonal submatrices are 1×1 matrices and $||x|| \equiv |x|$, then Theorem 2 We again remark that if the partitioning of (2.1) is such that all

satisfied (2.10) for at least one j, $1 \le j \le N$. each eigenvalue λ of an arbitrary $n \times n$ complex matrix A necessarily Inclusion regions for eigenvalues. In Theorem 2, we saw that

DEFINITION 3. For the partitioned $n \times n$ matrix A of (2.1), let the

Gerschgorin set G_i be the set of all complex numbers z such that

$$(\|(A_{j,j}-zI)^{-1}\|)^{-1} \leqq \sum_{\substack{k=1 \\ k \neq j}}^N \|A_{j,k}\| \ , \qquad \qquad 1 \leqq j \leqq N$$

is clear that each Gerschgorin set G_j is closed and bounded. Hence, so right side of (3.1) and independent of the vector norms used. Thus, from (2.5), we conclude that the Gerschgorin set G_j always contains the eigenvalues of $A_{j,j}$, independent to the magnitude of the Next, it

$$(3.2) \hspace{3.1em} G = \bigcup_{j=1}^N G_j \; .$$

each G_j . By Theorem 2, all the eigenvalues of A lie in G. Can any eigenvalue λ of A lie on the boundary of G? This can be answered trivially for the particular partitioning of (2.1'). In this case, the rightcase, Theorem 2 gives exact information about the eigenvalues of A. G is a finite point set consisting only of the eigenvalues of A. In this hand side of (3.1) is vacuously zero, and from (3.1), we see that the set Thus, we can speak of the boundary of G, as well as the boundary of

result of Taussky [11]. that A is block irreducible, which is the analogue of a well known It is interesting that Theorem 2 can be strengthened by the assumption

and let λ be an eigenvalue of A. If λ is a boundary point of G, then it is a boundary point of each set G_j , $1 \leq j \leq N$. THEOREM 3. Let the partitioned matrix A of (2.1) be block irreducible,

if $||W_j|| \le ||W_r|| = 1$, then as before Since λ is an eigenvalue of A, then $\sum_{j=1}^{N} A_{i,j}W_j = \lambda W_i$, and

(3.3)
$$(\|(A_{r,r} - \lambda I_r)^{-1}\|)^{-1} \le \sum_{\substack{j=1\\j \ne r}}^N \|A_{r,j}\| \| \|W_j\| \le \sum_{\substack{j=1\\j \ne r}}^N \|A_{r,j}\| .$$

to complete the proof of Theorem 1. irreducibility of A, the argument can be extended to every index j, But as λ is a boundary point of G, equality must hold throughout (3.3), showing that λ is a boundary point of G_r . Moreover, if $||A_{r,j}|| \neq 0$, then $1 \leq j \leq N$, which completes the proof. A similar argument can be applied this way, we conclude that λ is a boundary point of G_j . From the $||W_j||=1$, and we can repeat the argument with r replaced by j. In

proof, depending on a continuity argument, follows that given in [13, p. 287]. Another familiar result of Gerschgorin can also be generalized.

THEOREM 4. If the union $H=igcup_{j=1}^m G_{p_j}$, $1\leq p_j \leq N$, of m Gerschgorin

eigenvalues of A. sets is disjoint from partitioned matrix A of the remaining N (2.1), then H- m Gerschgorin sets containspreciselyfor the $\sum_{j=1}^m n_{p_j}$

illustration, consider the partitioned matrix inclusion regions for the eigenvalues of from the generalized form of Gerschgorin's Theorem 2. The previous example of the matrix of (2.1') indicated that sharper ಣ matrix A may To give another be obtained

$$(3.4) \qquad A = \begin{bmatrix} 4 & -2 & -1 & 0 \\ -2 & 4 & 0 & -1 \\ -1 & 0 & 4 & -2 \end{bmatrix} = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,3} \end{bmatrix}.$$

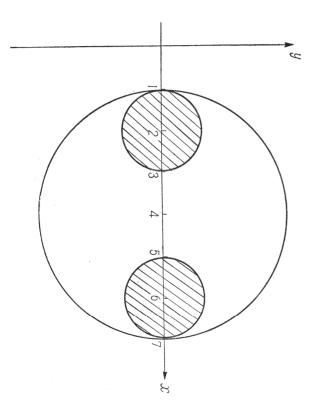
that $||A_{1,2}|| = ||A_{2,1}|| = 1$. Employing now the vector norm $||x||_2 \equiv$ On the other hand, direct computation shows $(\sum_i |x_i|^2)^{1/2}$, is apparent

$$(\|(A_{i,i}-zI_i)^{-1}\|^{-1}=\min\left\{|6-z|,|2-z|
ight\}, \qquad i=1,2$$
 .

By definition, the set G_1 then consists of the points z for which

$$|6-z| \le 1$$
, $|2-z| \le 1$,

since $G_2 = G_1$, as shown in the figure below. so that G_1 is itself the union of two disjoint circles. The same is true for G_2 , The usual Gerschgorin circles



significant improvements over the usual Gerschgorin circles in providing for the matrix A of (3.4) are all given by the single circle $|4-\lambda| \le 3$, which is a circle of radius 3, with center at z=4, as shown above. bounds for eigenvalues. For the matrix A of (3.4), its eigenvalues are From this figure, we conclude that the block Gerschgorin result can give

$$\lambda_{\scriptscriptstyle \rm I}=1$$
 , $\lambda_{\scriptscriptstyle \rm 2}=3$, $\lambda_{\scriptscriptstyle \rm 3}=5$, $\lambda_{\scriptscriptstyle \rm 4}=7$.

Note, again from the figure above, that Theorem 3 applies in this case. At this point, we remark that the previous example was such that

special case of each Gerschgorin set G_j consisted of the union of circles. This

set G_j is the union of n_j circles. diagonal submatrices $A_{j,j}$ are all normal. If the Euclidean vector norms $\|x\|_2$ are used for each subspace Ω_j , $1 \leq j \leq N$, then each Gerschgorin THEOREM 5. Let the partitioned matrix A of (2.1) be such that its

Proof. Let the eigenvalues of $A_{j,j}$ be σ_i , $1 \le l \le n_j$. Since $A_{j,j}$ is normal, we can write $(\|(A_{j,j}-zI_j)^{-1}\|)^{-1} = \min_l |\sigma_l-z|$, which, combined with Definition 3, completes the proof.

first example, the result of A. on inclusion regions for eigenvalues of $n \times n$ complex matrices. carries over. It is quite simple to obtain the block analogues of well known results Brauer [2] on ovals of Cassini easily

point sets $C_{i,j}$ defined by (2.1). Then, all the eigenvalues of A lie in the union of the [N(N-1)]/2Theorem 6. Let the $n \times n$ complex matrix A be partitioned as in

$$(3.5) \quad (\|(A_{i,i}-zI_i)^{-1}\|\cdot\|(A_{j,j}-zI_j)^{-1}\|)^{-1} \leq \left(\sum_{\substack{l=1\\l\neq i}}^{N}\|A_{i,l}\|\right)\left(\sum_{\substack{l=1\\l\neq j}}^{N}\|A_{j,l}\|\right)$$

where $1 \leq i, j \leq N$ and $i \neq j$. Moreover, if A is block irreducible, and λ is an eigenvalue of A not in the interior of $\bigcup_{i\neq j} C_{i,j}$, then λ is a

result by Ostrowski [9]. an illustration, we include the following known [4] generalization of a results using both row and column sums admit easy generalizations. leaves the eigenvalues of A invariant. Thus, rows sums can be replaced by column sums in the definition (2.4) of diagonal dominance, and many boundary point of each of the point sets $C_{i,j}$. Other obvious remarks can be made. Cle Clearly, replacing A by A^r

(2.1), and define THEOREM. 7. Let the $n \times n$ complex matrix A be partitioned as in

$$(3.6) \hspace{1cm} R_{j} \equiv \sum_{\substack{k=1 \\ k \neq j}}^{N} \|A_{j,k}\| \; ; \hspace{0.2cm} C_{j} \equiv \sum_{\substack{k=1 \\ k \neq j}}^{N} \|A_{k,j}\| \; , \hspace{1cm} 1 \leq j \leq N \; .$$

Then, for any α with $0 \le \alpha \le 1$, each eigenvalue λ of A satisfies

$$(\|(A_{j,j} - \lambda I_j)^{-1}\|)^{-1} \le R_j^{\alpha} C_j^{1-\alpha}$$

for at least one j, $1 \le j \le N$.

Also, the important result of Fan and Hoffman [3] carries over with

Let p > 1, and 1/p + 1/q = 1. If $\alpha > 0$ satisfies Let the $n \times n$ complex matrix A be partitioned as in

$$(3.8) \qquad \sum\limits_{i=1}^{N} \left\{ \frac{\left(\sum\limits_{j \neq i} \|A_{i,j}\|\right)^{q}}{\left(\sum\limits_{j \neq i} \|A_{i,j}\|^{p}\right)^{q/p}} \right\} \leqq \alpha^{q} (1 + \alpha^{q}) \; ,$$

(whenever 0/0 occurs on the left-hand side, we agree to put 0/0 = 0), then every eigenvalue λ of A satisfies at least one of the following relations:

$$(\|(A_{j,j}-\lambda I_j)^{-1}\|)^{-1} \leqq \alpha \left(\sum_{k=1 \atop k\neq j}^N \|A_{j,k}\|^p\right)^{1/p}, \qquad 1 \leqq j \leqq N.$$

all possible vector norms to produce optimum results. all the matrices $A_{i,j}$ of (2.1) are 1×1 matrices, these new inclusion regions now *depend* on the vector norms used. It seems reasonable, at partitioned, and this perhaps can be used to advantage. least theoretically, to minimize these inclusion regions by considering We wish to emphasize that, unlike the cases previously treated where great deal of flexibility in the manner in which the matrix A is Similarly, there

 $a_{i,i}$ $1 \leq i \leq n$, then the eigenvalues λ_j of A satisfy dominant in the usual sense of (2.4') with positive real diagonal entries [12], states that if an $n \times n$ matrix $A = (a_{i,j})$ is strictly diagonally Another generalization. Another result, due again to Taussky

$$(4.1) Re\lambda_j > 0 , 1 \leq j \leq n .$$

 x_i , let |x| denote the vector with components $|x_i|$. the following. which depends upon the use of absolute norms [1]. By this, we mean Based on our previous results, we now give a generalization of this result First, if x is a column vector with complex components

$$||x|| = |||x|||$$

for all vectors x, then the norm is an absolute norm. This is equivalent

⁴ Clearly, the l_p -norms of footnote 2 are absolute norms.

is a nonnegative real number, then [1] to the property that if $|y| \ge |x|$, i.e., each component of |y| - |x|

$$(4.2') \qquad \qquad ||\boldsymbol{y}|| \ge ||\boldsymbol{x}||.$$

and if B is nonsingular with $B^{-1} \equiv (r_{i,j})$ such that $r_{i,j} \geq 0$ for all $1 \neq j$, $j \leq m$, then B is said to be an M-matrix [8].

If λ is any eigenvalue of A, then $1 \leq j \leq N$, and the vector norms for each subspace Ω_j are absolute norms least one j). Further, assume that each submatrix $A_{i,j}$ is an M-matrix, and block diagonally dominant with strict inequality in (2.4) for at (2.1), and let A be block strictly diagonally dominant (or block irreducible Theorem 9. Let the $n \times n$ complex matrix A be partitioned as in

$$(4.3) Re\lambda > 0.$$

from the assumption that $A_{j,j}$ is an M-matrix that with $Rez \leq 0$. If $A_{j,j}^{-1} \equiv (r_{k,l})$, and $(A_{j,j} - zI_j)^{-1} \equiv (s_{k,l}(z))$, it follows [8] A is block strictly diagonally dominant. Let z be any complex number Proof. For simplicity, we shall consider again only the case where

$$|s_{k,l}(z)| \le r_{k,l} \; , \qquad \qquad 1 \le k, \, l \le n_j \; .$$

(4.2) and (4.2') that Next, with (4.4) and the assumption of absolute norms, it follows from

$$rac{\|(A_{j,j}-zI_j)^{-1}x\|}{\|x\|} \leq rac{\|A_{j,j}^{-1}|x|\|}{\||x\|\|},$$

so that from (2.3),

$$(||(A_{j,j}-zI_j)^{-1}||)^{-1}\geqq (||A_{j,j}^{-1}||)^{-1}$$
 .

if λ is an eigenvalue of A, then $Re\lambda > 0$, which completes the proof. In other words, for any z with $Rez \leq 0$, then the matrix A – to be block strictly diagonally dominant, and hence nonsingular. zI continues

BIBLIOGRAPHY

- F. L. Bauer, J. Stoer, and C. Witzgall, Absolute and monotonic norms, Numerische Math., 257 - 264.
- A. Brauer, Limits for the characteristic roots of a matrix II, Duke Math. J., 14 (1947).
- mination of Eigenvalues, edited by Olga Taussky, Nat. Bur. of Standards Appl. Math. Series of a matrix, Contributions to the Solution of Systems of Linear Equations and the **39** (1954), 117-130, Ky Fan and A. J. Hoffman, Lower bounds for the rank and location of the eigenvalues

- Rome, 1960, Birkhauser Verlag, (1961), 33-36. Treatment of Ordinary Differential Equations, Integral and Integro-Differential Equations, Miroslav Fiedler, Some estimates of spectra of matrices, Symposium on the Numerical
- Mat. Fyz. Časopis. Slovensk. Akad. Vied, 10 (1960), 148-166. Miroslay Fiedler and Vlastimil Pták, Some inequalities for the spectrum of a matrix
- (1960), 255–259. Frank Harary, On the consistency of precedence matrices, J. Assoc. Comput. Mach., 7
- **3** (1956), 314–324. Alston S. Householder, On the convergence of matrix iterations, J. Assoc. Comput. Mach,
- Math. Helv., 10 (1937), 69-96. A. M. Ostrowski, Über die Determinanten mit überwiegender Hauptdiagonale, Comment.
- 209-226.die Lokalisierung der charakteristischen Wurzeln von Matrizen, Compositio Math., 9 (1951), A. M. Ostrowski, Über das Nichtverschwinden einer Klasse von Determinanten und
- partitioned into blocks, Journal of Math. Anal. and Appl., 2 (1961), 161-209. 10. A. M. Ostrowski, On some metrical properties of operator matrices and matrices
- 1043-1044. 11. Olga Taussky, Bounds for characteristic roots of matrices, Duke Math. J., **15** (1948).
- 13. ———, Some topics concerning bounds for eigenvalues of finite matrices, Survey of Numerical Analysis, edited by John Todd, McGraw-Hill, (1962), 279-297. A recurring theorem on determinants, Amer. Math. Monthly, 56 (1949), 672-676.