ON LIMITS OF THE REAL CHARACTERISTIC ROOTS OF MATRICES WITH REAL ELEMENTS

GENE W. MEDLIN¹

A. Brauer [5]² (also see [4]) has obtained the following bound for matrices with arbitrary elements. Each characteristic root ω of the arbitrary matrix B must satisfy at least one of the $n(n-1)/2$ inequalities

$$(1) \quad |(Z - b_{\alpha\alpha})(Z - b_{\beta\beta}) - b_{\alpha\beta}b_{\beta\alpha}| \leq Q_{\alpha\beta}$$

where $\alpha, \beta = 1, 2, \dots, n; \alpha \neq \beta$ and

$$Q_{\alpha\beta} = \sum_{\nu=1; \nu \neq \alpha, \beta}^n (|b_{\alpha\beta}b_{\beta\nu}| + |b_{\beta\alpha}b_{\alpha\nu}| + |b_{\alpha\nu}b_{\beta\nu}|) \\ + \sum_{\nu=1; \nu < \mu; \nu, \mu \neq \alpha, \beta}^n |b_{\alpha\nu}b_{\beta\mu} + b_{\alpha\mu}b_{\beta\nu}|.$$

For the real roots of a matrix with real elements Brauer is able to replace the third expression in the parentheses by a number that is always less than or equal $\sum_{\nu=1; \nu \neq \alpha, \beta}^n |b_{\alpha\nu}b_{\beta\nu}|$. Moreover, by using the 2^n possible combinations of the signs of the components of the characteristic vector associated with the real root ω , he is able to get still lower bounds. Obviously, if n is very large the work involved in using the 2^n combinations of signs is prohibitive. The object of this paper is to show how better bounds than (1) may often be obtained for the real roots of matrices with real elements without using all 2^n possible combinations of signs.

First we prove a lemma. We say a function $u(x)$ is convex on the interval $a \leq x \leq b$ if for any three values of x , $a \leq x_1 < x_2 < x_3 \leq b$,

$$\begin{vmatrix} x_1 & u(x_1) & 1 \\ x_2 & u(x_2) & 1 \\ x_3 & u(x_3) & 1 \end{vmatrix} \geq 0.$$

This is equivalent to the alternate definition

Presented to the Society, November 18, 1955; received by the editors September 27, 1955.

¹ Research Participant, Mathematics Panel, Oak Ridge National Laboratory, (Summer, 1955) from Wake Forest College.

² Numbers in brackets refer to the bibliography at the end of the paper.

$$\frac{u(x_2) - u(x_1)}{x_2 - x_1} \leq \frac{u(x_3) - u(x_2)}{x_3 - x_2},$$

as can be shown by a few simple calculations.³

LEMMA [7]. If $u_i(x)$ ($i = 1, 2, 3, \dots, n$) is a finite set of functions, convex on the interval $a \leq x \leq b$, then $\sum_{i=1}^n u_i(x)$ is convex on the interval $a \leq x \leq b$.

PROOF.⁴ We use induction. Take any three values for x , ordered such that $a \leq x_1 < x_2 < x_3 \leq b$.

The lemma is true for $n = 1$.

Now we assume the lemma true for $n = k - 1$, and prove it true for $n = k$.

$$\begin{aligned} \Delta &= \begin{vmatrix} x_1 & \sum_{i=1}^k u_i(x_1) & 1 \\ x_2 & \sum_{i=1}^k u_i(x_2) & 1 \\ x_3 & \sum_{i=1}^k u_i(x_3) & 1 \end{vmatrix} = \begin{vmatrix} x_1 & \sum_{i=1}^{k-1} u_i(x_1) + u_k(x_1) & 1 \\ x_2 & \sum_{i=1}^{k-1} u_i(x_2) + u_k(x_2) & 1 \\ x_3 & \sum_{i=1}^{k-1} u_i(x_3) + u_k(x_3) & 1 \end{vmatrix} \\ &= \begin{vmatrix} x_1 & \sum_{i=1}^{k-1} u_i(x_1) & 1 \\ x_2 & \sum_{i=1}^{k-1} u_i(x_2) & 1 \\ x_3 & \sum_{i=1}^{k-1} u_i(x_3) & 1 \end{vmatrix} + \begin{vmatrix} x_1 & u_k(x_1) & 1 \\ x_2 & u_k(x_2) & 1 \\ x_3 & u_k(x_3) & 1 \end{vmatrix} \geq 0. \end{aligned}$$

This proves the lemma.

Before we prove our theorem let us introduce some notations. Let $A = (a_{\alpha\beta})$ be an n -square matrix with real elements.

Let

$$M_{\alpha\beta}^* = \sum_{\nu=1; \nu \neq \alpha, \beta}^n |a_{\alpha\beta} a_{\beta\nu} \epsilon + a_{\beta\alpha} a_{\alpha\nu}| \quad \text{for } \epsilon = 1,$$

³ Neither of these definitions is that used by Jensen [7]. His definition is that $u(x)$ is convex on the interval $a \leq x \leq b$ if, for every x_1 and x_2 , $u[(x_1 + x_2)/2] \leq [u(x_1) + u(x_2)]/2$. Geometrically, this means that the midpoint of any chord of $u(x)$ lies above or on the curve; hence the equivalence is evident, if $u(x)$ is continuous [8].

⁴ This proof was suggested by Professor R. C. F. Bartels, temporary employee (1955) at the Oak Ridge National Laboratory from the University of Michigan

$$m_{\alpha\beta}^* = \sum_{\nu=1; \nu \neq \alpha, \beta}^n |a_{\alpha\beta}a_{\beta\nu}\epsilon + a_{\beta\alpha}a_{\alpha\nu}| \quad \text{for } \epsilon = -1.$$

Let $M_{\alpha\beta} = \max [M_{\alpha\beta}^*, m_{\alpha\beta}^*]$.

In the expression $\sum_{\nu=1; \nu \neq \alpha, \beta}^n (a_{\alpha\nu}a_{\beta\nu})$ let $W_{\alpha\beta}^*$ denote the sum of the positive terms, $w_{\alpha\beta}^*$ the sum of the negative terms and $W_{\alpha\beta} = \max [W_{\alpha\beta}^*, |w_{\alpha\beta}^*|]$. Let

$$Q'_{\alpha\beta} = M_{\alpha\beta} + W_{\alpha\beta} + \sum_{\nu=1; \nu < \mu; \nu, \mu \neq \alpha, \beta}^n |a_{\alpha\nu}a_{\beta\mu} + a_{\alpha\mu}a_{\beta\nu}|.$$

THEOREM 1. *Each real characteristic root ω of A must satisfy at least one of the $n(n-1)/2$ inequalities $|(Z - a_{\alpha\alpha})(Z - a_{\beta\beta}) - a_{\alpha\beta}a_{\beta\alpha}| \leq Q'_{\alpha\beta}$ ($\alpha, \beta = 1, 2, \dots, n; \alpha \neq \beta$).*

PROOF. Since ω is a real characteristic root of A , there exists a real nontrivial solution $X = (x_1, x_2, x_3, \dots, x_n)$ to the system of linear equations

$$(2) \quad \sum_{\beta=1}^n a_{\alpha\beta}x_{\beta} = \omega x_{\alpha} \quad (\alpha = 1, 2, 3, \dots, n).$$

Assume that $1 = x_r \geq |x_s| \geq \max |x_{\nu}|$ ($\nu = 1, 2, \dots, n; \nu \neq r, s$). Setting $x_r = 1$ is no restriction since it is one form of normalization of the characteristic vector X' (transpose of X) associated with the characteristic root ω .

Consider the r th and s th equations of (2)

$$(3) \quad (\omega - a_{rr}) = \sum_{\nu=1; \nu \neq r}^n a_{r\nu}x_{\nu},$$

$$(4) \quad (\omega - a_{ss})x_s = \sum_{\mu=1; \mu \neq s}^n a_{s\mu}x_{\mu}.$$

If $x_s = 0$, then $x_{\nu} = 0$ for all $\nu \neq r$. Thus, from (3), we obtain $\omega = a_{rr}$ which proves the theorem for $x_s = 0$.

Assume $x_s \neq 0$. Multiply equations (3) and (4). We obtain

$$\begin{aligned} (\omega - a_{rr})(\omega - a_{ss})x_s &= \left(\sum_{\nu=1; \nu \neq r}^n a_{r\nu}x_{\nu} \right) \left(\sum_{\mu=1; \mu \neq s}^n a_{s\mu}x_{\mu} \right) \\ &= a_{rs}a_{sr}x_s + \sum_{\nu=1; \nu \neq r, s}^n (a_{rs}a_{s\nu}x_s + a_{sr}a_{r\nu})x_{\nu} \\ &\quad + \sum_{\nu=1; \nu \neq r, s}^n a_{r\nu}a_{s\nu}x_{\nu}^2 + \sum_{\nu=1; \nu < \mu; \nu, \mu \neq r, s}^n (a_{r\nu}a_{s\mu} + a_{r\mu}a_{s\nu})x_{\nu}x_{\mu}. \end{aligned}$$

Hence

$$(5) \quad \begin{aligned} & \left| (\omega - a_{rr})(\omega - a_{ss}) - a_{rs}a_{sr} \right| \leq \sum_{\nu=1; \nu \neq r, s}^n \left| a_{rs}a_{s\nu}x_s + a_{sr}a_{r\nu} \right| \left| \frac{x_\nu}{x_s} \right| \\ & + \left| \sum_{\nu=1; \nu \neq r, s}^n (a_{r\nu}a_{s\nu}) \frac{x_\nu^2}{x_s} \right| + \sum_{\nu=1; \nu < \mu; \nu, \mu \neq r, s}^n \left| a_{r\nu}a_{s\mu} + a_{r\mu}a_{s\nu} \right| \left| \frac{x_\nu x_\mu}{x_s} \right|. \end{aligned}$$

Consider the first term on the right side of (5).

$$(6) \quad \sum_{\nu=1; \nu \neq r, s}^n \left| a_{rs}a_{s\nu}x_s + a_{sr}a_{r\nu} \right| \left| \frac{x_\nu}{x_s} \right| \leq \sum_{\nu=1; \nu \neq r, s}^n \left| a_{rs}a_{s\nu}x_s + a_{sr}a_{r\nu} \right|.$$

Since on the interval $-1 \leq x \leq 1$ each term on the right of (6) is either a straight line segment in the upper half-plane (if $|(a_{sr}a_{r\nu})/(a_{rs}a_{s\nu})| \geq 1$) or two line segments lying in the upper half-plane with a common point on the x -axis at $x = -(a_{sr}a_{r\nu})/(a_{rs}a_{s\nu})$ (if $|(a_{sr}a_{r\nu})/(a_{rs}a_{s\nu})| < 1$), we have on this interval the sum of convex functions. (We use the interval -1 to $+1$ since $-1 \leq x_s \leq +1$.) By our lemma, $\sum_{\nu=1; \nu \neq r, s}^n |a_{rs}a_{s\nu}x_s + a_{sr}a_{r\nu}|$ is convex.

From our alternate definition of a complex function, it is evident that the maximum of a convex function on the closed interval $a \leq x \leq b$ is attained at an end point. Hence

$$(7) \quad \sum_{\nu=1; \nu \neq r, s}^n |a_{rs}a_{s\nu}x + a_{sr}a_{r\nu}| < M_{rs}.$$

Next consider the second term on the right side of (5). Now x_ν^2/x_s will take the sign of x_s . However the sum of the positive terms in this expression will be less than or equal W_{rs}^* if $x_s > 0$ or less than or equal $|w_{rs}^*|$ if $x_s < 0$. Similarly, the absolute value of the sum of the negative terms will be less than or equal $|w_{rs}^*|$ if $x_s > 0$ or less than or equal W_{rs}^* if $x_s < 0$. In any case

$$(8) \quad \left| \sum_{\nu=1; \nu \neq r, s}^n a_{r\nu}a_{s\nu} \frac{x_\nu^2}{x_s} \right| \leq W_{rs}.$$

Using (7) and (8) in (5), we get

$$\begin{aligned} & \left| (\omega - a_{rr})(\omega - a_{ss}) - a_{rs}a_{sr} \right| \\ & \leq M_{rs} + W_{rs} + \sum_{\nu=1; \nu < \mu; \nu, \mu \neq r, s}^n |a_{r\nu}a_{s\mu} + a_{r\mu}a_{s\nu}| \leq Q'_{rs}. \end{aligned}$$

This proves the theorem.

An analogous theorem holds for the columns of A .

If in Brauer's Theorems 12-17 [1], 18, 19 [2], and 22 [3] we require the matrix A to have real elements and in Theorems 18 and 19 require the polynomials to have real coefficients, then the theorems will remain correct if $P_\alpha P_\beta$ is replaced by $Q'_{\alpha\beta}$. The conditions for Brauer's Theorems 41 and 42 [4] may hold under this new result while they might not under previous results (other than the exception noted at the beginning).

In another recent paper [6], Brauer was able to obtain bounds for the ratios of the co-ordinates of the characteristic vectors of a matrix. He proved that

$$(9) \quad |x_\beta| \leq t_{r\beta}^{(\mu)} |x_r| \quad \text{where } \beta = 1, 2, \dots, n \text{ and } 0 < t_{r\beta}^{(\mu)} \leq 1.$$

The upper μ indicates that we get a sequence $t_{r\beta}^{(1)}, t_{r\beta}^{(2)}, \dots$ ($\beta = 1, 2, 3, \dots, n$).

Using (9) we may make the following improvements.

Let

$$M_{\alpha\beta}^{(\mu)*} = \sum_{\nu=1; \nu \neq \alpha, \beta}^n |a_{\alpha\beta} a_{\beta\nu} \epsilon + a_{\beta\alpha} a_{\alpha\nu}| t_{\alpha\nu}^{(\mu)} \quad \text{for } \epsilon = t_{\alpha\beta}^{(\mu)},$$

$$m_{\alpha\beta}^{(\mu)*} = \sum_{\nu=1; \nu \neq \alpha, \beta}^n |a_{\alpha\beta} a_{\beta\nu} \epsilon + a_{\beta\alpha} a_{\alpha\nu}| t_{\alpha\nu}^{(\mu)} \quad \text{for } \epsilon = -t_{\alpha\beta}^{(\mu)},$$

and

$$M_{\alpha\beta}^{(\mu)} = \max [M_{\alpha\beta}^{(\mu)*}, m_{\alpha\beta}^{(\mu)*}].$$

In the expression $\sum_{\nu=1; \nu \neq \alpha, \beta}^n a_{\alpha\nu} a_{\beta\nu} [t_{\alpha\nu}^{(\mu)}]^2$ let $W_{\alpha\beta}^{(\mu)*}$ denote the sum of the positive terms, $w_{\alpha\beta}^{(\mu)*}$ the sum of the negative terms and

$$W_{\alpha\beta}^{(\mu)} = \max [W_{\alpha\beta}^{(\mu)*}, |w_{\alpha\beta}^{(\mu)*}|].$$

Let

$$Q_{\alpha\beta}^{(\mu)'} = M_{\alpha\beta}^{(\mu)} + W_{\alpha\beta}^{(\mu)} + \sum_{\rho=1; \rho < \nu; \rho, \nu \neq \alpha, \beta}^n |a_{\alpha\rho} a_{\beta\nu} + a_{\alpha\nu} a_{\beta\rho}| t_{\alpha\rho}^{(\mu)} t_{\alpha\nu}^{(\mu)}.$$

In the same way that we proved Theorem 1, we may prove the following theorem.

THEOREM 2. *Each real characteristic root ω of A must satisfy at least one of the $n(n-1)/2$ inequalities*

$$|(Z - a_{\alpha\alpha})(Z - a_{\beta\beta}) - a_{\alpha\beta} a_{\beta\alpha}| \leq Q_{\alpha\beta}^{(\mu)'} \quad (\alpha, \beta = 1, 2, \dots, n; \alpha \neq \beta).$$

Again under the previously stated conditions, Brauer's Theorems 12-19 and 22 may have the $P_\alpha P_\beta$ replaced by the $Q_{\alpha\beta}^{(u)'}.$

BIBLIOGRAPHY

1. Alfred Brauer, *Limits for the characteristic roots of a matrix* II, Duke Math. J. vol. 14 (1947) pp. 21-26.
2. ———, *Limits for the characteristic roots of a matrix* III, Duke Math. J. vol. 15 (1948) pp. 871-977.
3. ———, *Limits for the characteristic roots of a matrix* IV: *Applications to stochastic matrices*, Duke Math. J. vol. 19 (1952) pp. 75-91.
4. ———, *Limits for the characteristic roots of a matrix* V, Duke Math. J. vol. 19 (1952) pp. 553-562.
5. ———, *Über die Lage der charakteristischen Wurzeln einer Matrix*, J. Reine Angew. Math. vol. 192 (1953) pp. 113-116.
6. ———, *Bounds for the ratios of the co-ordinates of the characteristic vectors of a matrix*, Proc. Nat. Acad. Sci. U.S.A. vol. 41 (1955) pp. 162-164.
7. J. L. W. V. Jensen, *Sur les fonctions convexes et les inégalités entre les valeurs moyennes*, Acta Math. vol. 30 (1906) pp. 175-193.
8. T. Bonnesen and W. Fenchel, *Theorie der Konvexen Körper*, New York, Chelsea, 1958.

WAKE FOREST COLLEGE