Brauer's ovals theorem

Canonical name	BrauersOvalsTheorem
Date of creation	2013-03-22 15:35:30
Last modified on	2013-03-22 15:35:30
Owner	Andrea Ambrosio (7332)
Last modified by	Andrea Ambrosio (7332)
Numerical id	15
Author	Andrea Ambrosio (7332)
Entry type	Algorithm
Classification	msc 15A42
Related topic	GershgorinsCircleTheorem

Let A be a square complex matrix, $R_i = \sum_{j \neq i} |a_{ij}| \quad 1 \leq i \leq n$. Let's consider the ovals of this kind: $O_{ij} = \{z \in \mathbb{C} : |z - a_{ii}|\,|z - a_{jj}| \leq R_i R_j\} \quad \forall i \neq j$. Such ovals are called *Cassini ovals*.

Theorem (A. Brauer): All the eigenvalues of A lie inside the union of these $\frac{n(n-1)}{2}$ ovals of Cassini:$\sigma(A) \subseteq \bigcup_{i \neq j} O_{ij}$.

Proof: Let $(\lambda, \mathbf{v})$ be an eigenvalue-eigenvector pair for A, and let v_p, v_q be the components of $\mathbf{v}$ with the two maximal absolute values, that is $|v_p| \geq |v_q| \geq |v_i| \quad \forall i \neq p$. (Note that $|v_p| \neq 0$, otherwise $\mathbf{v}$ should be all-zero, in contrast with eigenvector definition). We can also assume that $|v_q|$ is not zero, because otherwise $A\mathbf{v} = \lambda\mathbf{v}$ would imply $a_{pp} = \lambda$, which trivially verifies the thesis. Then, since $A\mathbf{v} = \lambda\mathbf{v}$, we have:

$(\lambda - a_{pp})v_p = \sum_{j=1, j \neq p}^{n} a_{pj} v_j$

and so

$|\lambda - a_{pp}|\,|v_p| = \left|\sum_{j=1, j \neq p}^{n} a_{pj} v_j\right| \leq \sum_{j=1, j \neq p}^{n} |a_{pj}|\,|v_j| \leq \sum_{j=1, j \neq p}^{n} |a_{pj}|\,|v_q| = R_p\,|v_q|$

that is

$|\lambda - a_{pp}| \leq R_p \frac{|v_q|}{|v_p|}$.

In the same way, we obtain:

$|\lambda - a_{qq}| \leq R_q \frac{|v_p|}{|v_q|}$.

Multiplying the two inequalities, the two fractional terms vanish, and we get:

$|\lambda - a_{pp}|\,|\lambda - a_{qq}| \leq R_p R_q$

which is the thesis.$\square$

Remarks:

1) Much like the Levy-Desplanques theorem states a sufficient condition, based on Gerschgorin circles, for non-singularity of a matrix, Brauer's theorem can be employed to state a similar sufficient condition; namely, the following result of Ostrowski holds:

Corollary: Let A be a $n \times n$ complex-valued matrix; if for all $i \neq j$ we have $|a_{ii}|\,|a_{jj}| > R_i R_j$, then A is non singular.

The proof is obvious, since, by Brauer's theorem, the above condition excludes the point $z = 0$ from the spectrum of A, implying this way $\det(A) \neq 0$.

2) Since both Gerschgorin's and Brauer's results rely upon the same $2n$ numbers, namely $\{a_{ii}\}_{i=1}^{n}$ and $\{R_i\}_{i=1}^{n}$, one may wonder if Brauer's result is stronger than Gerschgorin's one; actually, the answer is positive, as the following inclusion shows:

Corollary: Let $G(A) = \bigcup_{i=1}^{n} D_i(A)$ and $B(A) = \bigcup_{i \neq j}^{n} O_{ij}(A)$ be respectively Gershgorin and Brauer eigenvalues inclusion regions ($D_i(A)$ are the Gerschgorin circles and $O_{ij}(A)$ are the Brauer's Cassini ovals); then

$B(A) \subseteq G(A)$.

Proof: Let O_{ij} be one of the $n(n-1)/2$ ovals of Cassini for matrix A and be $z \in O_{ij}$. If $R_i = 0$ or $R_j = 0$, Brauer's theorem imply $z = a_{ii}$ or $z = a_{jj}$ respectively; but since both a_{ii} and a_{jj} belong to their respective Gerschgorin circles, we have $z \in (D_i \cup D_j)$. If both $R_i > 0$ and $R_j > 0$, then we can write:

$\frac{|z-a_{ii}|}{R_i} \cdot \frac{|z-a_{jj}|}{R_j} \leq 1$.

For the left-hand side to be not greater than 1, $\frac{|z-a_{ii}|}{R_i}$ or $\frac{|z-a_{jj}|}{R_j}$ must be not greater than 1, which in turn means $z \in D_i$ or $z \in D_j$, that is $z \in (D_i \cup D_j)$. This way, we proved that $O_{ij} \subseteq (D_i \cup D_j)$; now, we have:

$B(A) = \bigcup_{i \neq j} O_{ij} \subseteq \bigcup_{i=1}^{n} D_i = G(A)$.

3) It's obvious from definition that there are infinitely many matrices which generate the same ovals of Cassini: namely, let's define

$\Omega(A) = \{M \in \mathbf{C}^{n \times n} : m_{ii} = a_{ii}, R_i(M) = R_i(A)\}$

as the set of all matrices which share the same ovals of Cassini as A. Then, by Brauer's theorem, we have, for all $M \in \Omega$ matrices,

$$\sigma(M) \subseteq B(A),$$

and therefore, having defined $\sigma(\Omega) = \bigcup_{M \in \Omega} \sigma(M)$, we have

$$\sigma(\Omega) \subseteq B(A).$$

One may then ask how sharp this inclusion is, which, informally speaking, is equivalent to asking how "efficient" is the "use", by Brauer's theorem, of the 2n pieces of information $\{a_{ii}\}_{i=1}^{n}$ and $\{R_i\}_{i=1}^{n}$ in the construction of inclusion sets (if for example we found the inclusion to be very loose, that is $\sigma(\Omega)$ to be a very little subset of $B(A)$, we could conjecture that the knowledge of the 2n numbers used by Brauer's theorem should have led to a more precise bounding, since the spectra of all matrices which share these numbers lie in a much smaller region). It has been proven that actually

$$\sigma(\Omega) = B(A),$$

thus showing Brauer's ovals are *optimal ones* under this point of view.

References

[1] S. Gerschgorin, *Uber die Abgrenzung der Eigenwerte einer Matrix*, Isv. Akad. Nauk USSR Ser. Mat., 7 (1931), pp. 749-754

[2] A. Brauer, *Limits for the characteristic roots of a matrix II*, Duke Math. J. 14 (1947), pp. 21-26

[3] R. S. Varga and A. Krautstengl, *On Gersgorin-type problems and ovals of Cassini*, Electron. Trans. Numer. Anal., 8 (1999), pp. 15-20

[4] Richard S. Varga, *Gersgorin-type eigenvalue inclusion theorems and their sharpness*,Electronic Transactions on Numerical Analysis. Volume 12 (2001), pp. 113-133