Solving linear systems of equations

§3.7 Gerschgorin's Theorems

MA385 – Numerical Analysis 1

November 2018

There are some extra details posted as an "Appendix" to this section

The goal of this final section is to learn a technique for estimating eigenvalues of matrices.

The idea dates from 1931, and is a simple as it is useful. Although known to mathematicians in the USSR, the original paper was not widely read.

ИЗВЕСТИЯ АКАДЕМИИ НАУК СССР. 1931

BULLETIN DE L'ACADÉMIE DES SCIENCES DE L'URSS

Classe des sciences mathématiques et naturelles

Отделение математических и естественных наук

ÜBER DIE ABGRENZUNG DER EIGENWERTE EINER MATRIX

Von S. GERSCHGORIN

(Présenté par A. Krylov, membre de l'Académie des Sciences)

It received main-stream attention in the West following the work of Olga Taussky (*A recurring theorem on determinants*, American Mathematical Monthly, vol 56, p672–676. 1949.)

See also https://www.math.wisc.edu/hans/paper_archive/other_papers/hs057.pdf

(See Section 5.4 of Süli and Mayers).

Theorem 3.32 (Gerschgorin's First Theorem)

Given a matrix $A \in \mathbb{R}^{n\times n}$, define the n *Gerschgorin Discs*, $D_1, D_2, \dots, D_n$ as the discs in the complex plane where D_i has centre a_{ii} and radius r_i:

$$r_i = \sum_{j=1, j\neq i}^{n} |a_{ij}|.$$

So $D_i = \{z \in \mathbb{C} : |a_{ii} - z| \leq r_i\}$. All the eigenvalues of A are contained in the union of the Gerschgorin discs.

Proof.

The proof makes no assumption about A being symmetric, or the eigenvalues being real. However, if A is symmetric, then its eigenvalues are real and so the theorem can be simplified: the eigenvalues of A are contained in the union of the intervals $I_i = [a_{ii} - r_i, a_{ii} + r_i]$, for $i = 1, \ldots, n$.

Example 3.33

Let

$$A = \begin{pmatrix} 4 & -2 & 1 \\ -2 & -3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$$

Theorem 3.34 (Gerschgorin's Second Theorem)

Given a matrix $A \in \mathbb{R}^{n \times n}$, let the n Gerschgorin disks be as defined in Theorem 3.32. If k of discs are disjoint (have an empty intersection) from the others, their union contains k eigenvalues.

Proof: not covered in class. If interested, see the appendix, or the textbooks.

Example 3.35

Locate the regions contains the eigenvalues of

$$A = \begin{pmatrix} -3 & 1 & 2 \\ 1 & 4 & 0 \\ 2 & 0 & -6 \end{pmatrix}$$

(The eigenvalues are approximately -7.018, -2.130 and 4.144.)

Example 3.36

Use Gerschgorin's Theorems to find an upper and lower bound for the Singular Values of the matrix

$$A = \begin{pmatrix} 4 & -1 & 2 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{pmatrix}.$$

Hence give an upper bound for $\kappa_2(A)$.

Exercise 3.20

A real matrix $A = \{a_{i,j}\}$ is *Strictly Diagonally Dominant* if

$$|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{i,j}| \qquad \text{for } i = 1, \ldots, n.$$

Show that all strictly diagonally dominant matrices are nonsingular.

Exercise 3.21

Let

$$A = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & -3 \end{pmatrix}$$

Use Gerschgorin's theorems to give an upper bound for $\kappa_2(A)$.

Proof of Gerschgorin's First Theorem (Thm 3.32)

Let λ be an eigenvalues of A, so $A\boldsymbol{x} = \lambda x$ for the corresponding eigenvector $\boldsymbol{x}$. Suppose that x_i is the entry of $\boldsymbol{x}$ with largest absolute value. That is $|x_i| = \|\boldsymbol{x}\|_\infty$. Looking at the i^{th} entry of the vector $A\boldsymbol{x}$ we see that

$$(A\boldsymbol{x})_i = \lambda x_i \implies \sum_{j=1}^{n} a_{ij}x_j = \lambda x_i.$$

This can be rewritten as

$$a_{ii}x_i + \sum_{\substack{j=0 \\ j\neq i}}^{n} a_{ij}x_j = \lambda x_i,$$

which gives

$$(a_{ii} - \lambda)x_i = -\sum_{\substack{j=0 \\ j\neq i}}^{n} a_{ij}x_j$$

By the triangle inequality,

$$|a_{ii} - \lambda||x_i| = |\sum_{\substack{j=0 \\ j \neq i}}^{n} a_{ij} x_j| \leq \sum_{\substack{j=0 \\ j \neq i}}^{n} |a_{ij}||x_j| \leq |x_i| \sum_{\substack{j=0 \\ j \neq i}}^{n} |a_{ij}|,$$

since $|x_i| \geq |x_j|$ for all j. Dividing by $|x_i|$ gives

$$|a_{ii} - \lambda| \leq \sum_{\substack{j=0 \\ j \neq i}}^{n} |a_{ij}|,$$

as required.

Proof of Gerschgorin's 2nd Thm (Thm 3.34) We didn't do the proof in class, and you are not expected to know it. Here is a *sketch* of it.
Let $B(\varepsilon)$ be the matrix with entries

$$b_{ij} = \begin{cases} a_{ij} & i = j \\ \varepsilon a_{ij} & i \neq j. \end{cases}$$

So $B(1) = B$ and $B(0)$ is the diagonal matrix whose entries are the diagonal entries of A.

Each of the eigenvalues of $B(0)$ correspond to its diagonal entries and (obviously) coincide with the Gerschgorin discs of $B(0)$ – the centres of the Gerschgorin discs of A.

The eigenvalues of B are the zeros of the characteristic polynomial $\det(B(\varepsilon) - \lambda I)$ of B. Since the coefficients of this polynomial depend continuously on ε, so too do the eigenvalues.

Now as ε varies from 0 to 1, the eigenvalues of $B(\varepsilon)$ trace a path in the complex plane, and at the same time the radii of the Gerschgorin discs of A increase from 0 to the radii of the discs of A. If a particular eigenvalue was in a certain disc for $\varepsilon = 0$, the corresponding eigenvalue is in the corresponding disc for all ε.

Thus if one of the discs of A is disjoint from the others, it must contain an eigenvalue.

The same reasoning applies if k of the discs of A are disjoint from the others; their union must contain k eigenvalues.