Finding Eigenvalues

Up to this point, our main theoretical tools for finding eigenvalues without using $\det\{A - \lambda I\} = 0$ have been the trace and determinant formulas

$$\det\{A\} = \lambda_1 \lambda_2 \cdots \lambda_n, \quad \mathrm{Tr}\{A\} = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

plus the facts that

- the eigenvalues of a triangular matrix are the diagonal elements
- similar matrices $B = S^{-1}AS$ have the same eigenvalues
- the eigenvalues of a real symmetric matrix are real
- the eigenvalues of an orthogonal matrix have $|\lambda| = 1$.

The first fact can be generalized to block triangular matrices. If M is block upper triangular

$$M = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$$

the eigenvalues of M are the eigenvalues of A plus the eigenvalues of C since

$$0 = \det\{M - \lambda I\} = \det \begin{bmatrix} A - \lambda I & B \\ 0 & C - \lambda I \end{bmatrix} = \det\{A - \lambda I\} \det\{C - \lambda I\}.$$

Here each identity matrix I has the appropriate dimensions to match its partner. A similar statement holds for block lower triangular M. Thus for block triangular matrices, the eigenvalue problem can be broken up into smaller subproblems.

(A matrix M is *reducible* if it can be written in block triangular form by reordering rows and columns PMP^{-1}, where P is a permutation matrix.)

Gershgorin Circle Theorem

Also called the Gershgorin Disk Theorem.

Theorem statement and Examples 1 and 2 are based on LeVeque's *Finite Difference Methods for Ordinary & Partial Differential Equations.*

Theorem. Let $A \in \mathbb{C}^{n\times n}$ and let D_i be the closed disk in the complex plane centered at A_{ii} with radius given by the row sum $r_i = \sum_{j\neq i} |A_{ij}|$:

$$D_i = \{z \in \mathbb{C} : |z - A_{ii}| \leq r_i\} \equiv D(A_{ii}, r = r_i).$$

Then all the eigenvalues of A lie in the union of the disks D_i for $i = 1,\ \ldots,\ n$. If some set of k overlapping disks is disjoint from all the other disks, then exactly k eigenvalues lie in the union of these k disks.

Note the following:

- If a disk D_i is disjoint from all other disks, then it contains exactly one eigenvalue of A.
- If a disk D_i overlaps other disks, then it need not contain any eigenvalues, although the union of the overlapping disks contains the appropriate number.
- If A is real, A^T has the same eigenvalues as A and the theorem can also be applied to A^T (or equivalently the disk radii can be defined by summing the magnitudes of the off-diagonal elements of columns of A rather than rows).
- (If A is irreducible, a stronger version of the theorem states that an eigenvalue cannot lie on the boundary of a disk unless it lies on the boundary of every disk.)

Proof: $Ax = \lambda x$ implies

$$(\lambda - A_{ii})x_i = \sum_{j\neq i} A_{ij}x_j$$

$$|\lambda - A_{ii}| \leq \sum_{j\neq i} |A_{ij}| \frac{|x_j|}{|x_i|} \leq \sum_{j\neq i} |A_{ij}| = r_i$$

by choosing i such that $|x_i| = \max_j |x_j|$.

Examples

Here we assume all matrices are real.

1. 3×3 *example*

$$A = \begin{bmatrix} 5 & 0.6 & 0.1 \\ -1 & 6 & -0.1 \\ 1 & 0 & 2 \end{bmatrix}$$

The Gershgorin theorem applied to A implies that the eigenvalues lie within the union of $D(5, r = 0.7)$, $D(6, r = 1.1)$, and $D(2, r = 1)$. There is exactly one eigenvalue in $D(2, r = 1)$ and two eigenvalues in $D(5, r = 0.7) \cup D(6, r = 1.1)$.

All eigenvalues have real parts between 1 and 7.1 (and hence positive real parts, in particular). The eigenvalue in $D(2, r = 1)$ must be real, since complex eigenvalues must appear in conjugate pairs.

Applying the theorem to A^T gives a tighter bound on the single eigenvalue λ_1 near 2; λ_1 must lie within $D(2, r = 0.2)$, so $1.8 \le \lambda_1 \le 2.2$.

The actual eigenvalues of A are $\lambda_1 = 1.9639$ and $\lambda_\pm = 5.5180 \pm 0.6142i$.

2. *Second difference matrix*

The matrix $D^{(2)} = \text{tridiag}[1 \ -2 \ 1]$ is symmetric, so all its eigenvalues are real. By the Gershgorin theorem they must lie in the circle of radius 2 centered at -2: $-4 \le \lambda_j \le 0$. We know that $D^{(2)}$ is nonsingular, so all the eigenvalues of $D^{(2)}$ are negative. (Also since $D^{(2)}$ is irreducible and two circles have radii $= 1$, $-4 < \lambda_j < 0$, which proves $D^{(2)}$ is nonsingular.) The eigenvalues can actually be calculated:

$$\lambda_j = 2(\cos(j\pi h) - 1), \quad j = 1, \ \ldots, \ n$$

where $h = 1/(n+1)$, and thus are distributed between $-4 < \lambda_j < 0$.

The Jacobi iteration matrix is $B = \frac{1}{2}\text{tridiag}[1 \ 0 \ 1]$ with $-1 < \lambda_j < 1$ (since B is irreducible), and Jacobi converges since $\rho(B) < 1$.

3. *Absolute row sums* < 1

Suppose B has absolute row sums < 1. Then

$$r_i = |B_{i1}| + \cdots + |B_{i-1,1}| + |B_{i+1,1}| + \cdots + |B_{in}| < 1 - |B_{ii}|$$

and by the Gershgorin theorem all $|\lambda_i| < 1$ and $\rho(B) < 1$.

4. *Diagonally dominant matrix*

Suppose A is diagonally dominant:

$$r_i = |A_{i1}| + \cdots + |A_{i-1,1}| + |A_{i+1,1}| + \cdots + |A_{in}| < |A_{ii}|, \quad i = 1, \ \ldots, \ n.$$

Then by the Gershgorin theorem $\lambda_i \neq 0$ and A is invertible. For Jacobi iteration, $B = I - D^{-1}A$ has all absolute row sums < 1, so Jacobi converges. Gauss-Seidel and SOR are also guaranteed to converge.

In Examples 5–9, see how much information you can extract about the actual eigenvalues using just the Gershgorin Circle Theorem and (block) triangularity. The actual eigenvalues are given.

5. *From HW6*

$$A = \frac{1}{2}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$\lambda_1 = 0$, $\lambda_\pm = \pm\sqrt{2}/2$.

6. *From HW7*

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

$\lambda_\pm = (1 \pm \sqrt{5})/2$.

Last examples are from **Theory of Iterative Methods** *notes*

7. For

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$B_J = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}, \ \lambda_\pm = \pm\frac{1}{2}, \ \rho_J = \frac{1}{2}$$

$$B_{GS} = \begin{bmatrix} 0 & \frac{1}{2} \\ 0 & \frac{1}{4} \end{bmatrix}, \ \lambda_1 = 0, \ \lambda_2 = \frac{1}{4}, \ \rho_{GS} = \frac{1}{4}$$

8. *Example where Jacobi converges but Gauss-Seidel diverges*

$$A = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix}, \quad B_J = \begin{bmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix}$$

$\lambda_1 = \lambda_2 = \lambda_3 = 0$, $\rho_J = 0$, Jacobi converges.

$$B_{GS} = \begin{bmatrix} 0 & -2 & 2 \\ 0 & 2 & -3 \\ 0 & 0 & 2 \end{bmatrix}$$

$\lambda_1 = 0$, $\lambda_2 = \lambda_3 = 2$, $\rho_{GS} = 2$, Gauss-Seidel diverges.

9. Example where Jacobi diverges but Gauss-Seidel converges

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad B_J = \frac{1}{2} \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$$

$\lambda_1 = -1$, $\lambda_2 = \lambda_3 = 0.5$, $\rho_J = 1$, Jacobi diverges.

$$B_{GS} = \frac{1}{8} \begin{bmatrix} 0 & -4 & -4 \\ 0 & 2 & -2 \\ 0 & 1 & 3 \end{bmatrix}$$

$\lambda_1 = 0$, $\lambda_\pm = 0.3125 \pm 0.1654i$, $\rho_{GS} = 0.3536$, Gauss-Seidel converges.