Numerical Analysis Hilary Term 2023
Lecture 6: Matrix Eigenvalues

We now turn to eigenvalue problems $Ax = \lambda x$, where $A \in \mathbb{R}^{n\times n}$ or $A \in \mathbb{C}^{n\times n}$, $\lambda \in \mathbb{C}$, and $x(\neq 0) \in \mathbb{C}^n$. Recall that there are n eigenvalues in $\mathbb{C}$ (nonreal λ possible even if A is real). There are usually, but not always, n linearly independent eigenvectors (e.g. Jordan block $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ has only one eigenvector $[1, 0]^T$).

Background: An important result from analysis (not proved or examinable!), which will be useful.

Theorem. (Ostrowski) The eigenvalues of a matrix are continuously dependent on the entries. That is, suppose that $\{\lambda_i, i = 1, \ldots, n\}$ and $\{\mu_i, i = 1, \ldots, n\}$ are the eigenvalues of $A \in \mathbb{R}^{n\times n}$ and $A + B \in \mathbb{R}^{n\times n}$ respectively. Given any $\varepsilon > 0$, there is a $\delta > 0$ such that $|\lambda_i - \mu_i| < \varepsilon$ whenever $\max_{i,j} |b_{ij}| < \delta$, where $B = \{b_{ij}\}_{1\le i,j\le n}$.

Noteworthy properties related to eigenvalues:

- A has n eigenvalues (counting multiplicities), equal to the roots of the **characteristic polynomial** $p_A(\lambda) = \det(\lambda I - A)$.
- If $Ax_i = \lambda_i x_i$ for $i = 1, \ldots, n$ and x_i are linearly independent so that $[x_1, x_2, \ldots, x_n] =: X$ is nonsingular, then A has the **eigenvalue decomposition** $A = X\Lambda X^{-1}$. This usually, but not always, exist. The most general form is the Jordan canonical form (which we don't treat much in this course).
- Any square matrix has a **Schur decomposition** $A = QTQ^*$ where Q is unitary $QQ^* = Q^*Q = I_n$, and T triangular. The superscript $*$ denotes the (complex) conjugate transpose, $(Q^*)_{ij} = \overline{Q_{ji}}$.
- For a **normal matrix** s.t. $A^*A = AA^*$, the Schur decomposition shows T is diagonal (proof: examine diagonal elements of A^*A and AA^*), i.e., A can be diagonalized by a unitary similarity transformation: $A = Q\Lambda Q^*$, where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Most of the structured matrices we treat are normal, including symmetric ($\lambda \in \mathbb{R}$), orthogonal ($|\lambda| = 1$), and skew-symmetric ($\lambda \in i\mathbb{R}$).

Aim: estimate the eigenvalues of a matrix.

Theorem. Gerschgorin's theorem: Suppose that $A = \{a_{ij}\}_{1\le i,j\le n} \in \mathbb{R}^{n\times n}$, and λ is an eigenvalue of A. Then, λ lies in the union of the **Gerschgorin discs**

$$D_i = \left\{ z \in \mathbb{C} \,\middle|\, |a_{ii} - z| \le \sum_{\substack{j\neq i \\ j=1}}^{n} |a_{ij}| \right\}, \quad i = 1, \ldots, n.$$

Proof. If λ is an eigenvalue of $A \in \mathbb{R}^{n\times n}$, then there exists an eigenvector $x \in \mathbb{R}^n$ with $Ax = \lambda x$, $x \neq 0$, i.e.,

$$\sum_{j=1}^{n} a_{ij}x_j = \lambda x_i, \quad i = 1, \ldots, n.$$

Suppose that $|x_k| \geq |x_\ell|$, $\ell = 1, \ldots, n$, i.e.,

$$\text{“}x_k \text{ is the largest entry”}. \tag{1}$$

Then the kth row of $Ax = \lambda x$ gives $\displaystyle\sum_{j=1}^{n} a_{kj} x_j = \lambda x_k$, or

$$(a_{kk} - \lambda) x_k = - \sum_{\substack{j \neq k \\ j=1}}^{n} a_{kj} x_j.$$

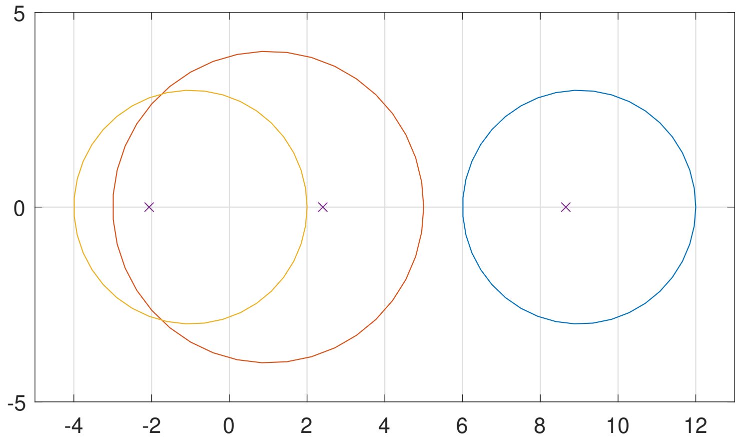
Dividing by x_k, (which, we know, is $\neq 0$) and taking absolute values,

$$|a_{kk} - \lambda| = \left| \sum_{\substack{j \neq k \\ j=1}}^{n} a_{kj} \frac{x_j}{x_k} \right| \leq \sum_{\substack{j \neq k \\ j=1}}^{n} |a_{kj}| \left| \frac{x_j}{x_k} \right| \leq \sum_{\substack{j \neq k \\ j=1}}^{n} |a_{kj}|$$

by (1). □

Example.

$$A = \begin{bmatrix} 9 & 1 & 2 \\ -3 & 1 & 1 \\ 1 & 2 & -1 \end{bmatrix}$$



With Matlab calculate `>> eig(A) = 8.6573, -2.0639, 2.4066`

Theorem. Gerschgorin's 2nd theorem: If any union of ℓ (say) discs is disjoint from the other discs, then it contains ℓ eigenvalues.

Proof. Consider $B(\theta) = \theta A + (1 - \theta) D$, where $D = \text{diag}(A)$, the diagonal matrix whose diagonal entries are those from A. As θ varies from 0 to 1, $B(\theta)$ has entries that vary continuously from $B(0) = D$ to $B(1) = A$. Hence the eigenvalues $\lambda(\theta)$ vary continuously by Ostrowski's theorem. The Gerschgorin discs of $B(0) = D$ are points (the diagonal entries), which are clearly the eigenvalues of D. As θ increases the Gerschgorin discs of $B(\theta)$ increase in radius about these same points as centres. Thus if $A = B(1)$ has a disjoint set of ℓ Gerschgorin discs by continuity of the eigenvalues it must contain exactly ℓ eigenvalues (as they can't jump!). □