
Numerical Analysis Hilary Term 2025

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 $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$).

Eigenvalues and roots of polynomials

Recall that the eigenvalues $\lambda_1, \dots, \lambda_n$ (repeated for multiplicity) are the roots of the degree n characteristic polynomial $\det(A - \lambda I)$. So, we could compute the eigenvalues of A by computing the roots of the characteristic polynomial. Conversely, we can reformulate a polynomial root-finding problem into an eigenvalue problem. Let $p(x) = \sum_{i=0}^n c_i x^i$ be a degree n polynomial with real coefficients and consider the following matrix, which is called the **companion matrix** (the blank elements are all 0) for p :

$$C = \begin{bmatrix} -\frac{c_{n-1}}{c_n} & -\frac{c_{n-2}}{c_n} & \dots & -\frac{c_1}{c_n} & -\frac{c_0}{c_n} \\ 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}$$

Note that if $p(\lambda) = 0$ for some $\lambda \in \mathbb{C}$, then

$$C \begin{bmatrix} \lambda^{n-1} \\ \lambda^{n-2} \\ \vdots \\ \lambda \\ 1 \end{bmatrix} = \begin{bmatrix} -\sum_{i=0}^{n-1} \frac{c_i}{c_n} \lambda^i \\ \lambda^{n-1} \\ \lambda^{n-2} \\ \vdots \\ \lambda \end{bmatrix} = \begin{bmatrix} \lambda^n \\ \lambda^{n-1} \\ \lambda^{n-2} \\ \vdots \\ \lambda \end{bmatrix} = \lambda \begin{bmatrix} \lambda^{n-1} \\ \lambda^{n-2} \\ \vdots \\ \lambda \\ 1 \end{bmatrix},$$

and so λ is an eigenvalue of C whose corresponding eigenvector is $x = [\lambda^{n-1}, \lambda^{n-2}, \dots, \lambda, 1]^\top$. Indeed, one can show that the characteristic polynomial $\det(\lambda I - C) = p(\lambda)/c_n$ (nonexchangeable), so the eigenvalues of C are precisely the roots of p , counting multiplicities. Thus, any eigenvalue problem can be expressed as a polynomial-root finding problem, and conversely, any polynomial-root finding problem can be expressed as an eigenvalue problem.

One idea to compute the eigenvalues of A is to compute the roots of the characteristic polynomial. However, if $n \geq 5$, then there is no closed form solution for the roots. Thus, any algorithm we use to compute eigenvalues must be *iterative*.

Direct vs. Iterative Methods:

Methods such as LU or QR factorisations and solving $Ax = b$ using them are *direct*: they compute a certain number of operations and then finish with “the answer”. For eigenvalues, we need **iterative** methods, which

-
- construct a sequence and
 - truncate that sequence “after convergence”.

Iterative methods are typically concerned with fast convergence rate (rather than operation count).

Properties related to eigenvalues and eigenvalue estimates

Before we introduce several iterative algorithms, we first review some theoretical properties of eigenvalues.

- If $Ax_i = \lambda_i x_i$ for $i = 1, \dots, n$ and x_i are linearly independent so that $[x_1, x_2, \dots, x_n] =: X$ 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 in this course).
- Any square matrix has a **Schur decomposition** $A = QTQ^*$ where Q is unitary $QQ^* = Q^*Q = I_n$, and T triangular whose diagonal entries are the eigenvalues of A . 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 $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$. Most of the structured matrices we treat are normal, including symmetric ($\lambda \in \mathbb{R}$), orthogonal ($|\lambda| = 1$), and skew-symmetric ($\lambda \in i\mathbb{R}$).

We can also (roughly) estimate the eigenvalues of a matrix $A \in \mathbb{R}^{n \times n}$ using $\|A\|_2$.

Lemma. If λ is an eigenvalue of A , then $|\lambda| \leq \|A\|_2$.

Proof. If $Ax = \lambda x$ with $\|x\|_2 = 1$, then

$$|\lambda| = \|Ax\|_2 = \|A\|_2 \|x\|_2 = \|A\|_2,$$

which completes the proof. □

As a second attempt, we will use two theorems from Gerschgorin that provide theoretical bounds for every eigenvalue λ (rather than just $|\lambda|$), although these bounds are usually not precise enough in practice.

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

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

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.,

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

Let $k := \arg \max_{\ell=1,\dots,n} |x_\ell|$ so that “ x_k is the largest entry.” Then, the k th row of $Ax = \lambda x$ gives $\sum_{j=1}^n a_{kj}x_j = \lambda x_k$, or

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

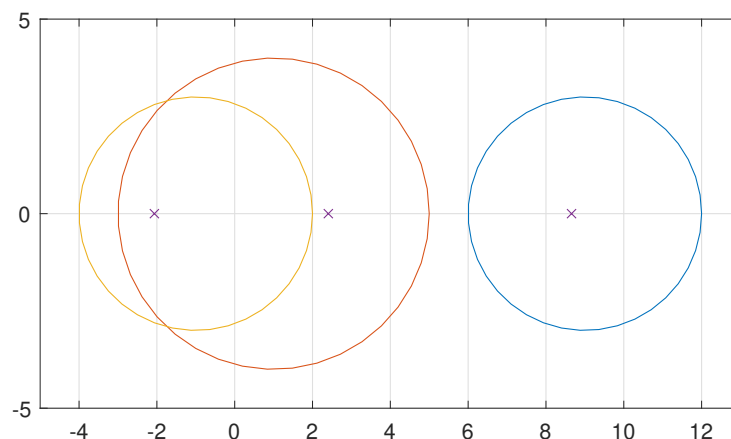
Dividing by x_k , (which, we know, is $\neq 0$ as $x = 0$ otherwise) and taking absolute values, we obtain

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

which completes the proof. □

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`

We can improve upon Gerschgorin’s first theorem, using an important result from analysis (not proved or examinable!):

Theorem. (Ostrowski) The eigenvalues of a matrix are continuously dependent on the entries. That is, let $A, B \in \mathbb{R}^{n \times n}$ and suppose that $\{\lambda_i, i = 1, \dots, n\}$ and $\{\mu_i, i = 1, \dots, n\}$ are the eigenvalues of A and $A + B$ 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 $B = \{b_{ij}\}_{1 \leq i,j \leq n}$.

Theorem. Gerschgorin’s 2nd theorem: Suppose that for some $\ell \in \{1, \dots, n\}$, the union of ℓ Gerschgorin discs is disjoint from the remaining $(n - \ell)$ discs. Then, this union 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!). \square

Gerschgorin's theorems are particularly useful when either (i) A is close to diagonal or (ii) when one can use a similarity transform to obtain a nearly diagonal matrix. Recall that a matrix A is similar to B if there is a nonsingular matrix P for which $A = P^{-1}BP$. Similar matrices have the same eigenvalues, since if $A = P^{-1}BP$,

$$0 = \det(A - \lambda I) = \det(P^{-1}(B - \lambda I)P) = \det(P^{-1})\det(P)\det(B - \lambda I),$$

so $\det(A - \lambda I) = 0$ if, and only if, $\det(B - \lambda I) = 0$.

Example. Consider

$$A = \begin{bmatrix} 1 & \epsilon & \epsilon \\ \epsilon & 4 & 1 \\ \epsilon & 1 & 5 \end{bmatrix}.$$

For $\epsilon \ll 1$, we would expect one eigenvalue of A to be close to 1 since the off-diagonal terms in the first row and column are close to zero. Indeed, the Gerschgorin discs are

$$D_1 = \{z \in \mathbb{C} : |z - 1| < 2\epsilon\}, \quad D_2 = \{z \in \mathbb{C} : |z - 4| < 1 + \epsilon\}, \quad D_3 = \{z \in \mathbb{C} : |z - 5| < 1 + \epsilon\}.$$

For $\epsilon \ll 1$, the first disc is disjoint from the rest, so Gerschgorin's second theorem shows that one eigenvalue λ satisfies $|\lambda - 1| \leq 2\epsilon$.

Using a similarity transformation, we can obtain a sharper bound. Let

$$D = \begin{bmatrix} \epsilon & & \\ & 1 & \\ & & 1 \end{bmatrix}.$$

Then, $B = DAD^{-1}$ has the same eigenvalues of A , but

$$B = \begin{bmatrix} 1 & \epsilon^2 & \epsilon^2 \\ 1 & 4 & 1 \\ 1 & 1 & 5 \end{bmatrix},$$

and so the Gerschgorin discs are now

$$D_1 = \{z \in \mathbb{C} : |z - 1| < 2\epsilon^2\}, \quad D_2 = \{z \in \mathbb{C} : |z - 4| < 2\}, \quad D_3 = \{z \in \mathbb{C} : |z - 5| < 2\}.$$

Again, for $\epsilon \ll 1$, the first disc is disjoint from the rest, so an eigenvalue satisfies $|\lambda - 1| < 2\epsilon^2$, which is a sharper bound as $\epsilon \rightarrow 0$ than the one we obtained using the Gerschgorin discs of A .

Estimating Extremal Eigenvalues: The Power Method

We now turn to a simple method for calculating a single (largest) eigenvalue of a square matrix A (and its associated eigenvector).

Notation: in iterative methods, x_k usually means the vector x at the k th iteration (rather than k th entry of vector x). Some sources use x^k or $x^{(k)}$ instead.

Algorithm 1 Power Method

Require: $A \in \mathbb{R}^{n \times n}$ and $x_0 \in \mathbb{R}^n$

for $k = 1, 2, \dots$ **do**

$y_k = Ax_k$

$x_{k+1} = y_k / \|y_k\|_2$

end for

Algorithm 1 is the **Power Method** or **Power Iteration**, and computes unit vectors in the direction of $x_0, Ax_0, A^2x_0, A^3x_0, \dots, A^kx_0$.

Lemma. Suppose that $A \in \mathbb{R}^{n \times n}$ is diagonalizable and let $\{\lambda_i, v_i\}_{i=1}^n$ denote the eigenvalues and eigenvectors of A , where

- the eigenvalues are sorted in decreasing magnitude $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$ and
- the eigenvectors have unit norm $\|v_i\|_2 = 1$ for $i = 1, \dots, n$.

If $|\lambda_1| > |\lambda_2|$, then for any “generic” x_0 , the iterates x_k of the power method satisfy either

$$\|x_k - \operatorname{sgn}(\lambda_1)^k v_1\|_2 = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right) \text{ or } \|x_k + \operatorname{sgn}(\lambda_1)^k v_1\|_2 = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right) \text{ as } k \rightarrow \infty,$$

and

$$|\lambda| - \|y_k\|_2 = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right) \text{ as } k \rightarrow \infty,$$

where

$$\operatorname{sgn}(\lambda_1) := \begin{cases} 1 & \text{if } \lambda > 0, \\ 0 & \text{if } \lambda = 0, \\ -1 & \text{if } \lambda < 0. \end{cases}$$

That is, x_k converges to $\pm v_1$ (meaning that x_k may alternate signs every iteration if $\lambda_1 < 0$) and $\|y_k\|_2$ converges to $|\lambda|$ at a rate $|\lambda_2/\lambda_1|$. (The assumption that x_0 is “generic” means that we have $x_0 = \sum_{i=1}^n \alpha_i v_i$ with $\alpha_1 \neq 0$.) We can get the sign of λ by computing the ratio $(Ax_k)_1/(x_k)_1$.

Proof.

1. Note that $x_1 = Ax_0/\|Ax_0\|_2$ and

$$x_k = \frac{A^k x_0}{\|A^k x_0\|_2} \implies x_{k+1} = \frac{Ax_k}{\|Ax_k\|_2} = \frac{A^{k+1} x_0}{\|A^{k+1} x_0\|_2},$$

so by induction, $x_k = A^k x_0 / \|A^k x_0\|_2$. As a result, $y_k = A^{k+1} x_0 / \|A^k x_0\|_2$.

2. By assumption, we have

$$x_0 = \sum_{i=1}^n \alpha_i v_i$$

for some $\alpha_i \in \mathbb{R}$, $i = 1, 2, \dots, n$ with $\alpha_1 \neq 0$, and so

$$A^k x_0 = A^k \sum_{i=1}^n \alpha_i v_i = \sum_{i=1}^n \alpha_i A^k v_i = \sum_{i=1}^n \alpha_i \lambda_i^k v_i = \lambda_1^k \left[\alpha_1 v_1 + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i \right]$$

since $Av_i = \lambda_i v_i$. Moreover,

$$\|A^k x_0\|_2 = |\lambda_1|^k \left\| \alpha_1 v_1 + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i \right\|_2 =: |\lambda_1|^k \beta_k,$$

and so

$$x_k = \frac{\text{sgn}(\lambda_1)^k}{\beta_k} \left(\alpha_1 v_1 + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i \right).$$

At this point, we can see that $\beta_k \rightarrow |\alpha_1|$ since $\|v_1\|_2 = 1$ and, at least formally, that “ $x_k \rightarrow \pm v_1$ ” at a rate $|\lambda_2/\lambda_1|$. Then, we also have

$$\|y_k\|_2 = \|Ax_k\|_2 \rightarrow \|Av_1\|_2 = |\lambda| \quad \text{as } k \rightarrow \infty.$$

The remainder of the proof is just providing the technical details and nonexaminable.

3. We then have

$$\|x_k - \text{sgn}(\alpha_1) \text{sgn}(\lambda_1)^k v_1\|_2 = \left\| \alpha_1 \left(\frac{1}{\beta_k} - \frac{1}{|\alpha_1|} \right) v_1 + \frac{1}{\beta_k} \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i \right\|_2,$$

where we used that $\text{sgn}(\alpha_1) = \alpha_1/|\alpha_1|$. Applying the triangular inequality then gives

$$\begin{aligned} \|x_k - \text{sgn}(\alpha_1) \text{sgn}(\lambda_1)^k v_1\|_2 &\leq \left\| \alpha_1 \left(\frac{1}{\beta_k} - \frac{1}{|\alpha_1|} \right) v_1 \right\|_2 + \left\| \frac{1}{\beta_k} \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i \right\|_2 \\ &\leq \left| \frac{|\alpha_1| - \beta_k}{\beta_k} \right| + \frac{1}{\beta_k} \left| \frac{\lambda_2}{\lambda_1} \right|^k \sum_{i=2}^n |\alpha_i|, \end{aligned}$$

where we used that $\|v_i\|_2 = 1$ for $i = \{2, \dots, n\}$. Now, we can use the reverse triangle inequality to obtain

$$||\alpha_1| - \beta_k| = ||\alpha_1 v_1\|_2 - \beta_k| \leq \left\| \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i \right\|_2 \leq \left| \frac{\lambda_2}{\lambda_1} \right|^k \sum_{i=2}^n |\alpha_i|.$$

Thus,

$$\|x_k - \text{sgn}(\alpha_1) \text{sgn}(\lambda_1)^k v_1\|_2 \leq \frac{2}{\beta_k} \left| \frac{\lambda_2}{\lambda_1} \right|^k \sum_{i=2}^n |\alpha_i|.$$

The final step is to bound β_k from below. Using the reverse triangle inequality again, we have

$$\beta_k \geq \|\alpha_1 v_1\|_2 - \sum_{i=2}^n \alpha_i \left\| \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i \right\|_2 \geq |\alpha_1| - \left| \frac{\lambda_2}{\lambda_1} \right|^k \sum_{i=2}^n |\alpha_i|.$$

The second term converges to zero monotonically as $k \rightarrow \infty$, so for k sufficiently large, $\beta_k \geq |\alpha_1 v_1|/2$. Thus, for k sufficiently large, we have

$$\|x_k - \operatorname{sgn}(\alpha_1) \operatorname{sgn}(\lambda_1)^k v_1\|_2 \leq \frac{4}{|\alpha_1|} \left| \frac{\lambda_2}{\lambda_1} \right|^k \sum_{i=2}^n |\alpha_i|.$$

For the convergence of $\|y_k\|_2$, we can again use the reverse triangle inequality

$$\begin{aligned} \left| \|y_k\|_2 - |\lambda| \right| &= \left| \|Ax_k\|_2 - \|\operatorname{sgn}(\alpha_1) \operatorname{sgn}(\lambda_1)^k A v_1\|_2 \right| \\ &\leq \|A(x_k - \operatorname{sgn}(\alpha_1) \operatorname{sgn}(\lambda_1)^k v_1)\|_2 \\ &\leq \|A\|_2 \|x_k - \operatorname{sgn}(\alpha_1) \operatorname{sgn}(\lambda_1)^k v_1\|_2, \end{aligned}$$

which completes the proof. □

Note: it is unlikely but possible for a chosen vector x_0 that $\alpha_1 = 0$, but rounding errors in the computation generally introduce a small component in v_1 , so that in practice this is not a concern!

Inverse Power Method: If we apply the power method to A^{-1} , then the resulting method is called the **inverse power method**. Since the largest eigenvalue of A^{-1} is $1/\lambda_n$, where $\{\lambda_1, \dots, \lambda_n\}$ are the eigenvalues of A , the inverse power method can be used to compute the (inverse of) the smallest eigenvalue of A .