Chapter 4 Computation of Eigenvalues and Eigenvectors

Eigenvalues and eigenvectors are fundamental concepts in linear algebra with numerous applications in science and engineering. This course section will cover the localization of eigenvalues and the power method for their computation.

4.1 Localization of Eigenvalues

Definition

Eigenvalues are scalar values associated with a square matrix A that satisfy the Eq. (48):

$$A\mathbf{x} = \boldsymbol{\lambda}\mathbf{x} \tag{48}$$

Where $\boldsymbol{\lambda}$ is an eigenvalue and x is the corresponding eigenvector.

The equation states that when the matrix A acts on the vector x, the output is a scaled version of x. In other words, A transforms x by merely stretching or compressing it without changing its direction.

Example;

$A = \begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix}$ and x=$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$

$Ax = \begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 3 \times \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

Thus, the eigenvalue $\boldsymbol{\lambda} = \mathbf{3}$ and the associated eigenvector is $x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

4.1.1 Finding Eigenvalues: Analytical calculation

To find the eigenvalues of a matrix A, we rearrange the equation (48) into the following form:

$$\mathrm{Ax} - \lambda\mathrm{x} = 0 \tag{49}$$

This can be rewritten as:

$$(A - \lambda I)x = 0 \tag{50}$$

where I is the identity matrix of the same size A. For non-trivial solutions (i.e., x≠0), the determinant of $(A - \lambda I)$must be zero:

$$\boldsymbol{det}(\boldsymbol{A} - \boldsymbol{\lambda I}) = \mathbf{0} \tag{51}$$

he equation $\boldsymbol{det}(\boldsymbol{A} - \boldsymbol{\lambda I}) = 0$ is known as the characteristic equation of the matrix A The solutions to this polynomial equation give the eigenvalues λ1,λ2,…,λn .

Example

Consider the matrix:

$$A = \begin{pmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{pmatrix}$$

To find the Eigenvalues and eigenvectors, we first need to compute the eigenvalues of the matrix A.

1) Characteristic Polynomial: The eigenvalues are found by solving the characteristic equation given by: $det(A - \lambda I) = 0$.

   Where I is the identity matrix. For our matrix A:

$$A = \begin{pmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{pmatrix} \quad and\ I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$A - \lambda I = \begin{pmatrix} 10-\lambda & 0 & 0 \\ 1 & -3-\lambda & -7 \\ 0 & 2 & 6-\lambda \end{pmatrix}$$

2) Determinant Calculation**:**

$$det(A - \lambda I) = \det \begin{pmatrix} 10-\lambda & 0 & 0 \\ 1 & -3-\lambda & -7 \\ 0 & 2 & 6-\lambda \end{pmatrix} = (10-\lambda)[(-3-\lambda)(6-\lambda)+14]$$
$$= (\lambda - 10)[\lambda^2 - 3\lambda - 4] = (\lambda - 10)(\lambda + 1)(\lambda - 4)$$

3) Solving the Quadratic Equation: $(\lambda - 10)(\lambda + 1)(\lambda - 4) = 0$

   This gives us: $\lambda_1 = 10$, $\lambda_2 = -1\ and \quad \lambda_3 = 4$

The Eigenvalues are $\lambda_1 = 10$, $\quad \lambda_2 = -1\ and \quad \lambda_3 = 4$

4) Calculating the eigenvectors:

   Now, we compute the eigenvectors:

For $\lambda_1 = 10$

$$Ax = \lambda_1 x \rightarrow \begin{pmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 10\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \rightarrow \begin{cases} 10x_1 = 10x_1 \\ x_1 - 3x_2 - 7x_3 = 10x_2 \\ 2x_2 + 6x_3 = 10x_3 \end{cases}$$
$$\rightarrow \begin{cases} x_1 = 1 \\ x_2 = 2/33 \\ x_3 = 1/33 \end{cases} \rightarrow X_1 = \begin{pmatrix} 1 \\ 2/33 \\ 1/33 \end{pmatrix}$$

For $\lambda_2 = -1$

$$Ax = \lambda_2 x \rightarrow \begin{pmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = -1 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \rightarrow \begin{cases} 10x_1 = -x_1 \\ x_1 - 3x_2 - 7x_3 = -x_2 \\ 2x_2 + 6x_3 = -x_3 \end{cases}$$

$$\rightarrow \begin{cases} x_1 = 0 \\ x_2 = 1 \\ x_3 = -2/7 \end{cases} \rightarrow X_2 = \begin{pmatrix} 0 \\ 1 \\ -2/7 \end{pmatrix}$$

For $\lambda_3 = 4$

$$Ax = \lambda_3 x \rightarrow \begin{pmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 4 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \rightarrow \begin{cases} 10x_1 = 4x_1 \\ x_1 - 3x_2 - 7x_3 = 4x_2 \\ 2x_2 + 6x_3 = 4x_3 \end{cases}$$

$$\rightarrow \begin{cases} x_1 = 0 \\ x_2 = -1 \\ x_3 = 1 \end{cases} \rightarrow X_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

The Eigenvectors are $X_1 = \begin{pmatrix} 1 \\ 2/33 \\ 1/33 \end{pmatrix}, X_2 = \begin{pmatrix} 0 \\ 1 \\ -2/7 \end{pmatrix}$ ***and*** $X_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$

Unfortunately, this approach is generally not recommended, as there's no reliable and quick way to find the roots of a polynomial with a degree higher than 4. Thus, we investigate numerical methods for solving the eigenvalue problem

4.1.2 Localization Techniques

Localization methods help to determine the approximate locations of eigenvalues without computing them directly. Some common techniques include:

4.1.2.1 Gershgorin Circle Theorem

The Gershgorin Circle Theorem states that every eigenvalue of a matrix $A = [a_{ij}]$ lies within at least one of the Gershgorin disks defined by Eq. (52):

$$D_i = \{z \in C: [z - a_{ij}] \leq \sum_{j \neq i} |a_{ij}|\} \quad (52)$$

This means that for each row i we can draw a circle in the complex plane centered at a_{ij} with a radius equal to the sum of the absolute values of the other entries in the row.

The Gerschgorin theorem states that all the eigenvalues of a matrix A belong to the union of n disks. If the Gerschgorin disks are all disjoint, then each one contains exactly one eigenvalue.

Notations

1) All the eigenvalues of A are located within the union of the disks.
2) The i^{th} disk is defined with center a_{ij} and radius r_i is given by Eq. (53):

$$r_i = \sum_{j \neq i}^{n} |a_{ij}| \quad (53)$$

The Gerschgorin theorem provides a valuable way to localize eigenvalues of a matrix. Each Gerschgorin disk can be visualized in the complex plane, centered at the diagonal entry $a_{ij}$of the matrix and extending outwards by the radius r_i.

For example, if you have a matrix:

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

the disks are calculated as follows:

1. **For i=1**

   - Center: : $b_{11} = 1$

   The Gerschgorin disks according to the rows:

   Radius: $r_1 = |0| + |-3| = 3$

   Disk: $D_1 \rightarrow (1,3)$

   The Gerschgorin disks according to the columns

   Radius: $r_1 = |2| + |1| = 3$

   Disk: $D_1 \rightarrow (1,3)$

2. **For i=2**:

   Center: $b_{22} = 3$

   The Gerschgorin disks according to the rows:

   Radius: : $r_2 = |2| + |1| = 3$

   Disk: $D_2 \rightarrow (3,3)$

   The Gerschgorin disks according to the columns:

   Radius: : $r_2 = |0| + |0| = 0$

   Disk: $D_2 \rightarrow (3,0)$

3. **For i=3**

   Center: $b_{33} = -2$

   The Gerschgorin disks according to the rows:

Radius: : $r_3 = |1| + |0| = 1$

Disk: $D_3 \to (-2,1)$

The Gerschgorin disks according to the columns:

Radius: : $r_3 = |-3| + |1| = 4$

Disk: $D_3 \to (-2,4)$

Knowing that B and B^T share the same eigenvalues, we choose the smallest radius for each Gerschgorin disk. This provides the following localization of the spectrum.

$$D_1 \to (1,3)\ , \qquad D_2 \to (3,0) \qquad and \quad D_3 \to (-2,1)$$

Figure 4 shows the localization of the real eigenvalues and complex eigenvalues.

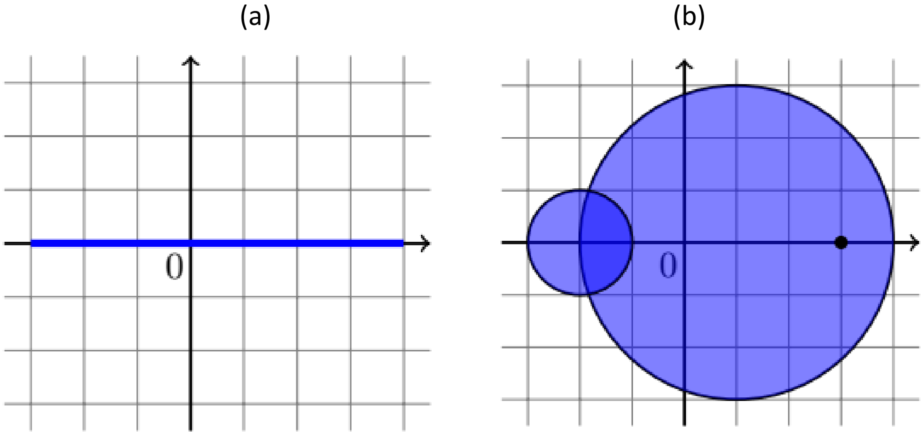

Figure.4 *Localization of the real eigenvalues (a) and complex eigenvalues (b).*

By employing analytical calculations, we were able to accurately determine the eigenvalues, which were then used to verify localization.

Determinant Calculation**:**

$$\mathrm{Det}(B - \lambda I) = \det\begin{bmatrix} 1-\lambda & 0 & -3 \\ 2 & 3-\lambda & 1 \\ 1 & 0 & -2-\lambda \end{bmatrix}$$

$$=(1-\lambda)\big[\big((3-\lambda)(-2-\lambda)-0\big)\big] + 0 - [3(3-\lambda)\big]$$

$$= (1-\lambda)[(\lambda-3)(\lambda-2)] + 3(\lambda-3)]$$

$$=(\lambda-3)[-(\lambda-1)(\lambda-2)+3]$$

$$=(\lambda-3)[-(\lambda^2-3\lambda+2)+3]$$

=-$(\lambda-3)[\lambda^2+\lambda+1]$

Solving the Quadratic Equation:

$$-(\lambda-3)[\lambda^2+\lambda+1]=0 \rightarrow \lambda-3=0 \quad or \quad \lambda^2+\lambda+1=0$$

The solutions are $\lambda_1=3,\ \lambda_2=-\frac{1}{2}+\frac{\sqrt{3}}{2}i, and\ \lambda_3=-\frac{1}{2}-\frac{\sqrt{3}}{2}$i . Figure 4 (b) confirms the localization of the eigenvalues..

Exercise: Consider the following matrix

A= $\begin{pmatrix} 30 & 1 & 2 & 3 \\ 4 & 15 & -4 & -2 \\ -1 & 0 & 3 & 5 \\ -3 & 5 & 0 & -1 \end{pmatrix}$

1) Calculate the Gershgorin disks.
2) Draw the Gershgorin disks in the complex plane.
3) Determine the possible localization of the eigenvalues of matrix A based on the Gershgorin disks.

In addition to the Gershgorin Circle Theorem, there are several other localization methods for determining eigenvalues. ***Interval Analysis*** is one such method; for instance, using ***Sturm's theorem***, we can ascertain the number of eigenvalues within a specified interval by analyzing the roots of the characteristic polynomial. This helps in refining the intervals that contain the eigenvalues.

Furthermore, various ***numerical methods*** can be utilized to obtain eigenvalues with greater precision after initial localization. Techniques like the ***QR algorithm*** offer a reliable way to compute eigenvalues more accurately, enhancing our understanding of the spectral properties of matrices.

4.2 Power Method

The Power Method is one of the simplest techniques for calculating the eigenvalues of a matrix A. This iterative method is particularly useful for finding the eigenvalue with the largest absolute value, known as the dominant eigenvalue, along with its corresponding eigenvector.

Theorem

Let A be a square matrix of size $n \times n$ that has N eigenvalues. The eigenvalues are defined as the values $\lambda_1, \lambda_2\ \ldots,\ \lambda_n$ that satisfy the characteristic equation Eq. (54):

$$\left|\boldsymbol{\lambda_1}\right| > |\boldsymbol{\lambda_2}| \geq\ \ldots \geq |\boldsymbol{\lambda_n}| \tag{54}$$

An eigenvalue λ_1 of a matrix A is said to be **dominant** if its absolute value is greater than the absolute values of all other eigenvalues of A.

Let $X^{(0)}$be a suitably chosen vector, then the sequences of vectors in Eq. (55)

$\left\{X^{(k)} = \left[x_1^{(k)}, x_2^{(k)} \ldots, x_n^{(k)}\right]\right\}$ (55)

And the sequence of scalars C_k generated in Eq. (56) where Any vector X in R^n can be expressed as:$\sum_{i=1}^{n} c_i x_i$.

$x^{(1)} = Ax^{(0)}$=$\sum_{i=1}^{n} c_i A x_i$=$\sum_{i=1}^{n} c_i \lambda_i \, x_i$ (56)

$$x^{(k)} = A^k x^{(0)} = \sum_{i=1}^{n} c_i (\lambda_i)^k \, x_i = \lambda_1^k \left[c_i x_i + c_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k x_2 + \cdots + c_2 \left(\frac{\lambda_n}{\lambda_1}\right)^k x_n\right]$$

Converges respectively to the dominant eigenvector $\boldsymbol{v_1}$ and the eigenvalue λ_1.

NB. The dominant eigenvalue plays a crucial role in various numerical methods and applications, particularly in iterative methods for solving systems of linear equations or finding eigenvalues themselves. The dominant eigenvalue significantly impacts various aspects of numerical analysis and system behavior. In iterative methods like the Power Method, it governs the convergence rate, allowing the method to quickly approach the dominant eigenvalue and its associated eigenvector, while other eigenvalues influence the process to a lesser extent. Furthermore, in dynamic systems, the magnitude of the dominant eigenvalue serves as an indicator of stability: if it is less than 1, the system is stable, whereas if it exceeds 1, the system becomes unstable. Additionally, the dominant eigenvalue provides valuable insights into the long-term behavior of processes modeled by matrices, including applications in population dynamics, Markov chains, and iterative algorithms.

Iterative Process

The method proceeds as follows:

1) **Initialization**: Choose an initial vector $X^{(0)}$ (often randomly).
2) **Iteration**: For k=0,1,2,…

   $x^{(1)} = Ax^{(0)}$=$\sum_{i=1}^{n} c_i A x_i$=$\sum_{i=1}^{n} c_i \lambda_i \, x_i$

   Here, c_i are the coefficients corresponding to the eigenvalues λi and eigenvectors x_i.

3) **Normalization**: To avoid overflow or underflow, normalize $x^{(1)}$:

   $$Y^{(k+1)} = c^{(k+1)} X^{(k)}$$

   where $c^{(k+1)} = \max_{1 \le i \le N}\left\{\left|{x_i}^{(k)}\right|\right\}$, This ensures that the components of the vector $Y^{(k+1)}$ are scaled to maintain numerical stability.

4) **Update the Iterative Vector**:

   $$X^{(k+1)} = \frac{1}{c^{(k+1)}} X^{(k)}$$

   Here, ${c_i}^{(k+1)} = {x_j}^{(k)}$ is the coefficient that contributes to the eigenvector approximation.

5) **Convergence Check**: Estimate the dominant eigenvalue using the Rayleigh quotient:

$$\lambda \approx \frac{X^{(k)T} A X^{(k)}}{X^{(k)T} X^{(k)}}$$

Repeat until the change in $X^{(k)}$ is smaller than a predefined tolerance.

Example:

Applying the Power Method to Matrix A

Given the matrix A= $\begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix}$ and the initial vector $x^{(0)}$= $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

we will apply the Power Method to find the dominant eigenvalue and its corresponding eigenvector.

Iterative Process

1. **First Iteration**:

$$\mathrm{A}x^{(0)}=\begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}=\begin{bmatrix} 10 \\ 1 \\ 0 \end{bmatrix}$$

Normalization: Calculate the maximum element for normalization:

$$c^{(1)} = \max_{1 \le i \le N}\{|x_i{}^{(1)}|\} = 10$$

Normalize $x^{(1)}$:

$$Y^{(1)} = \frac{X^{(1)}}{c^{(1)}} = \frac{1}{10}\begin{bmatrix} 10 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0,1 \\ 0 \end{bmatrix}$$

Update**:**

$$X^{(1)} = Y^{(1)}.c^{(1)} = \begin{bmatrix} 1 \\ 0,1 \\ 0 \end{bmatrix}.10 = \begin{bmatrix} 10 \\ 1 \\ 0 \end{bmatrix}$$

2. **Second Iteration**:

$$x^{(2)} =\mathrm{A}x^{(1)}=\begin{bmatrix} 10 & 0 & 0 \\ 1 & -3 & -7 \\ 0 & 2 & 6 \end{bmatrix}\begin{bmatrix} 10 \\ 1 \\ 0 \end{bmatrix}=\begin{bmatrix} 100 \\ -20 \\ 12 \end{bmatrix}$$

$$c^{(2)} = \max_{1 \le i \le N}\{|x_i{}^{(1)}|\} = 100$$

Normalize $x^{(2)}$:

$$Y^{(2)} = \frac{X^{(2)}}{c^{(2)}} = \frac{1}{100}\begin{bmatrix}100\\-20\\12\end{bmatrix} = \begin{bmatrix}1\\-0,2\\0.12\end{bmatrix}$$

Update**:**

$$X^{(2)} = Y^{(2)}.c^{(2)} = \begin{bmatrix}1\\-0,2\\0.12\end{bmatrix}.100 = \begin{bmatrix}100\\-20\\12\end{bmatrix}$$

3. Third Iteration:

$$x^{(3)} = \mathrm{A}x^{(2)} = \begin{bmatrix}10 & 0 & 0\\1 & -3 & -7\\0 & 2 & 6\end{bmatrix}\begin{bmatrix}100\\-20\\12\end{bmatrix} = \begin{bmatrix}1000\\-140\\72\end{bmatrix}$$

$$c^{(3)} = \max_{1\le i\le N}\{|x_i{}^{(1)}|\} = 1000$$

Normalize $x^{(2)}$:

$$Y^{(3)} = \frac{X^{(3)}}{c^{(3)}} = \frac{1}{1000}\begin{bmatrix}1000\\-140\\72\end{bmatrix} = \begin{bmatrix}1\\-0,14\\0.072\end{bmatrix}$$

Update**:**

$$X^{(2)} = Y^{(2)}.c^{(2)} = \begin{bmatrix}1\\-0,14\\0.072\end{bmatrix}.1000 = \begin{bmatrix}1000\\-140\\72\end{bmatrix}$$

Rayleigh Quotient

Finally, we can estimate the dominant eigenvalue using the Rayleigh quotient:

$$\lambda \approx \frac{X^{(2)T}AX^{(2)}}{X^{(2)T}X^{(k2)}} = \frac{[100 \quad -20 \quad 12]\begin{bmatrix}1000\\-140\\72\end{bmatrix}}{[100 \quad -20 \quad 12]\begin{bmatrix}100\\-20\\12\end{bmatrix}} = \frac{103664}{10444} \approx 9.91$$

The Power Method has shown that the dominant eigenvalue of matrix A is approximately λ≈9.91, with an associated eigenvector that converges through the iterative process.

Exercise:

Consider the following matri x and Initial Vector

A= $\begin{pmatrix} 5 & 4 & 2 \\ 2 & 3 & 1 \\ 1 & 2 & 3 \end{pmatrix}$ and $X^{(0)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

1) Apply the Power Method to find the dominant eigenvalue and its corresponding eigenvector of matrix A.
2) Perform two iterations of the Power Method and calculate the Rayleigh quotient to estimate the dominant eigenvalue.