Unit IV: The Power Method

Introduction

Given a square matrix A of size $n \times n$, the Power Method can be used to approximate the largest eigenvalue $\lambda_{\max}$ (in absolute value) and its corresponding eigenvector $\mathbf{v}$.

Example 1: Finding the Dominant Eigenvalue and Eigenvector

Let's apply the Power Method to a simple 3x3 matrix:

$$A = \begin{pmatrix} -3 & 0 & 0 \\ -4 & 6 & 2 \\ 16 & -15 & -5 \end{pmatrix}$$

Step 1: Initialization

Start with an initial guess for the eigenvector $\mathbf{x}^{(0)}$:

$$\mathbf{x}^{(0)} = \begin{pmatrix} 1 \\ 0.5 \\ 0.25 \end{pmatrix}$$

Step 2: Iteration

Apply the Power Method for a few iterations to see the convergence:

Iteration 1:

$$\begin{aligned} \mathbf{y}^{(1)} &= A\mathbf{x}^{(0)} = \begin{pmatrix} -3 & 0 & 0 \\ -4 & 6 & 2 \\ 16 & -15 & -5 \end{pmatrix} \begin{pmatrix} 1 \\ 0.5 \\ 0.25 \end{pmatrix} = \begin{pmatrix} 3 \\ -0.5 \\ 7.25 \end{pmatrix} \\ \mu_1 &= 7.25 \\ \mathbf{x}^{(1)} &= \frac{\mathbf{y}^{(1)}}{\mu_1} = \begin{pmatrix} 0.413793 \\ -0.068966 \\ 1 \end{pmatrix} \end{aligned}$$

Iteration 2:

$$
\begin{aligned}
\mathbf{y}^{(2)} &= A\mathbf{x}^{(1)} = \begin{pmatrix} 1.241379 \\ -0.068966 \\ 2.655172 \end{pmatrix} \\
\mu_2 &= 2.655172 \\
\mathbf{x}^{(2)} &= \frac{\mathbf{y}^{(2)}}{\mu_2} = \begin{pmatrix} 0.467532 \\ -0.025974 \\ 1.000000 \end{pmatrix}
\end{aligned}
$$

Iteration 3:

$$
\begin{aligned}
\mathbf{y}^{(3)} &= A\mathbf{x}^{(2)} = \begin{pmatrix} 1.402597 \\ -0.025974 \\ 2.870130 \end{pmatrix} \\
\mu_3 &= 2.870130 \\
\mathbf{x}^{(3)} &= \frac{\mathbf{y}^{(3)}}{\mu_3} = \begin{pmatrix} 0.488688 \\ -0.009050 \\ 1.000000 \end{pmatrix}
\end{aligned}
$$

Algorithm

The Power Method involves the following steps:

1. **Initialization**: Start with an arbitrary non-zero vector $\mathbf{x}^{(0)}$ (usually chosen randomly or with all components equal).

2. **Iteration**: For $k = 0, 1, 2, \ldots$, repeat the following steps until convergence:

$$\mathbf{y}^{(k+1)} = A\mathbf{x}^{(k)}$$

$$\mathbf{x}^{(k+1)} = \frac{\mathbf{y}^{(k+1)}}{\mu_k} \quad (\mu_k = y_i^k \text{ with } ||\mathbf{y}^k||_\infty = |y_i^k|)$$

   or

$$\mathbf{x}^{(k+1)} = \frac{\mathbf{y}^{(k+1)}}{\|\mathbf{y}^{(k+1)}\|}$$

3. **Convergence**: The sequence $\mathbf{x}^{(k)}$ converges to the eigenvector corresponding to the dominant eigenvalue $\lambda_{\max}$. The eigenvalue can be approximated as:

$$\lambda_{\max} \approx \frac{\mathbf{x}^{(k+1)} \cdot A\mathbf{x}^{(k)}}{\mathbf{x}^{(k)} \cdot \mathbf{x}^{(k)}}$$

Example 2: Finding the Dominant Eigenvalue and Eigenvector

Let's apply the Power Method to a simple 2x2 matrix:

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

Step 1: Initialization

Start with an initial guess for the eigenvector $\mathbf{x}^{(0)}$:

$$\mathbf{x}^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Step 2: Iteration

Apply the Power Method for a few iterations to see the convergence:

Iteration 1:

$$\mathbf{y}^{(1)} = A\mathbf{x}^{(0)} = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$$

Normalize $\mathbf{y}^{(1)}$ to obtain $\mathbf{x}^{(1)}$:

$$\mathbf{x}^{(1)} = \frac{\mathbf{y}^{(1)}}{\|\mathbf{y}^{(1)}\|} = \frac{1}{\sqrt{5^2 + 5^2}} \begin{pmatrix} 5 \\ 5 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0.7071 \\ 0.7071 \end{pmatrix}$$

Iteration 2:

$$\mathbf{y}^{(2)} = A\mathbf{x}^{(1)} = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 0.7071 \\ 0.7071 \end{pmatrix} = \begin{pmatrix} 3.5355 \\ 3.5355 \end{pmatrix}$$

Normalize $\mathbf{y}^{(2)}$ to obtain $\mathbf{x}^{(2)}$:

$$\mathbf{x}^{(2)} = \frac{\mathbf{y}^{(2)}}{\|\mathbf{y}^{(2)}\|} = \begin{pmatrix} 0.7071 \\ 0.7071 \end{pmatrix}$$

Iteration 3:

$$\mathbf{y}^{(3)} = A\mathbf{x}^{(2)} = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 0.7071 \\ 0.7071 \end{pmatrix} = \begin{pmatrix} 3.5355 \\ 3.5355 \end{pmatrix}$$

Normalize $\mathbf{y}^{(3)}$ to obtain $\mathbf{x}^{(3)}$:

$$\mathbf{x}^{(3)} = \frac{\mathbf{y}^{(3)}}{\|\mathbf{y}^{(3)}\|} = \begin{pmatrix} 0.7071 \\ 0.7071 \end{pmatrix}$$

Convergence

After a few iterations, the vector $\mathbf{x}^{(k)}$ stabilizes, indicating convergence to the dominant eigenvector. The corresponding eigenvalue can be approximated as:

$$\lambda_{\max} \approx \frac{\mathbf{x}^{(k+1)} \cdot A\mathbf{x}^{(k)}}{\mathbf{x}^{(k)} \cdot \mathbf{x}^{(k)}}$$

In this case:

$$\lambda_{\max} \approx \frac{\begin{pmatrix} 0.7071 & 0.7071 \end{pmatrix} \begin{pmatrix} 3.5355 \\ 3.5355 \end{pmatrix}}{\begin{pmatrix} 0.7071 \\ 0.7071 \end{pmatrix} \cdot \begin{pmatrix} 0.7071 \\ 0.7071 \end{pmatrix}} = \frac{4.9997}{0.9999} \approx 5$$

So, the dominant eigenvalue is approximately 5, and the corresponding eigenvector is $\mathbf{x} \approx \begin{pmatrix} 0.7071 \\ 0.7071 \end{pmatrix}$.

Result

After several iterations, the vector $\mathbf{x}^{(k)}$ will converge to the eigenvector corresponding to the dominant eigenvalue. The approximate eigenvalue can be computed as shown in the algorithm.

Convergence

The Power Method converges if the matrix A has a dominant eigenvalue that is strictly greater in magnitude than the other eigenvalues. The rate of convergence depends on the ratio of the dominant eigenvalue to the second largest eigenvalue.

Limitations

- The Power Method may fail or converge slowly if there is no clear dominant eigenvalue.
- If A has complex eigenvalues, modifications to the method are required.

Inverse Power Method

Introduction

The Inverse Power Method is an iterative technique used to find the eigenvalue of a matrix A that is closest to a given shift μ, as well as the corresponding eigenvector. This method is particularly effective for finding the smallest eigenvalue when $\mu = 0$.

Algorithm

The Inverse Power Method involves the following steps:

1. **Initialization**: Choose an initial guess for the eigenvector $\mathbf{x}^{(0)}$ and a shift μ.

2. **Iteration**: For $k = 0, 1, 2, \ldots$, repeat the following steps until convergence:
$$\mathbf{y}^{(k+1)} = (A - \mu I)^{-1}\mathbf{x}^{(k)}$$
$$\mathbf{x}^{(k+1)} = \frac{\mathbf{y}^{(k+1)}}{\|\mathbf{y}^{(k+1)}\|}$$

3. **Eigenvalue Approximation**: After convergence, the eigenvalue λ closest to μ can be approximated by:
$$\lambda \approx \mu + \frac{1}{\mathbf{x}^{(k+1)} \cdot \mathbf{y}^{(k+1)}}$$

Example

Consider the matrix A:
$$A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$$

Suppose we want to find the eigenvalue closest to $\mu = 0$ using the Inverse Power Method.

1. **Initialization**:
$$\mathbf{x}^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mu = 0$$

2. **Iteration** (for a few steps):

    Iteration 1:
$$\mathbf{y}^{(1)} = (A - 0 \cdot I)^{-1}\mathbf{x}^{(0)} = A^{-1}\mathbf{x}^{(0)}$$

    Compute A^{-1}:
$$A^{-1} = \frac{1}{4 \cdot 3 - 2 \cdot 1}\begin{pmatrix} 3 & -1 \\ -2 & 4 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 3 & -1 \\ -2 & 4 \end{pmatrix}$$

    Then,
$$\mathbf{y}^{(1)} = \frac{1}{10}\begin{pmatrix} 3 & -1 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0.2 \\ 0.2 \end{pmatrix}$$

    Normalize $\mathbf{y}^{(1)}$ to obtain $\mathbf{x}^{(1)}$:
$$\mathbf{x}^{(1)} = \frac{\mathbf{y}^{(1)}}{\|\mathbf{y}^{(1)}\|} = \begin{pmatrix} 0.7071 \\ 0.7071 \end{pmatrix}$$

Iteration 2:

$$\mathbf{y}^{(2)} = A^{-1}\mathbf{x}^{(1)} = \frac{1}{10}\begin{pmatrix} 3 & -1 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} 0.7071 \\ 0.7071 \end{pmatrix} = \begin{pmatrix} 0.1414 \\ 0.1414 \end{pmatrix}$$

Normalize $\mathbf{y}^{(2)}$ to obtain $\mathbf{x}^{(2)}$:

$$\mathbf{x}^{(2)} = \begin{pmatrix} 0.7071 \\ 0.7071 \end{pmatrix}$$

Convergence

The Inverse Power Method converges rapidly when μ is close to an eigenvalue of A. The rate of convergence depends on how close the shift μ is to the eigenvalue of interest.

Applications

- Finding the smallest eigenvalue of A (by setting $\mu = 0$).
- Refining eigenvalue approximations obtained by other methods.
- Solving eigenvalue problems for matrices where direct methods are computationally expensive.

Convergence of the Power Method

Theorem

Let A be an $n \times n$ diagonalizable matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Suppose that $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$, i.e., λ_1 is the dominant eigenvalue. Then, for almost every initial vector $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$ with $c_1 \neq 0$, the sequence generated by the power method:

$$\mathbf{x}_k = \frac{A\mathbf{x}_{k-1}}{\mu_k}, \quad k = 1, 2, \ldots \; (\mu_k = (A\mathbf{x}_{k-1})_i \text{ with } ||Ax_{k-1}||_\infty = |(Ax_{k-1})_i|)$$

satisfies:

1. $\mathbf{x}_k$ converges to an eigenvector corresponding to λ_1. Moreover it also satisfies
$$||\mathbf{x}_k - K\mathbf{v}_1|| \leq C_1 \left(\frac{\lambda_2}{\lambda_1}\right)^k,$$
2. The Rayleigh quotient $\mu_k = \frac{\mathbf{x}_k^T A\mathbf{x}_k}{\mathbf{x}_k^T \mathbf{x}_k}$ converges to λ_1. Furthermore, it also holds:
$$|\lambda_1 - \mu_k| \leq C_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k.$$

Proof

Since A is diagonalizable, its eigenvectors form a basis for $\mathbb{R}^n$. Thus, the initial vector $\mathbf{x}_0$ can be written as:

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n,$$

with $c_1 \neq 0$ (which holds for almost every $\mathbf{x}_0$).

Then,

$$A^k\mathbf{x}_0 = c_1\lambda_1^k\mathbf{v}_1 + c_2\lambda_2^k\mathbf{v}_2 + \cdots + c_n\lambda_n^k\mathbf{v}_n.$$

Factoring out λ_1^k:

$$A^k\mathbf{x}_0 = \lambda_1^k\left[c_1\mathbf{v}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{v}_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{v}_n\right].$$

Since $|\lambda_1| > |\lambda_j|$ for $j \geq 2$, we have $\left(\frac{\lambda_j}{\lambda_1}\right)^k \to 0$ as $k \to \infty$. Therefore,

$$A^k\mathbf{x}_0 \approx \lambda_1^k c_1\mathbf{v}_1 \quad \text{for large } k.$$

So, $A^k\mathbf{x}_0$ converges to a multiple of $\mathbf{v}_1$.

Now, the power method iteration is:

$$\mathbf{x}_k = \frac{A\mathbf{x}_{k-1}}{\mu_k} = \frac{A^2\mathbf{x}_{k-2}}{\mu_k\mu_{k-1}} = \cdots = \frac{A^k\mathbf{x}_0}{\mu_k\mu_{k-1}\cdots\mu_1}.$$

Hence, we have

$$\mathbf{x}_k = m_k A^k\mathbf{x}_0,$$

where $m_k = 1/(\mu_1\mu_2\cdots\mu_k)$. But, $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n,\ c_1 \neq 0$. Therefore

$$\mathbf{x}_k = m_k\lambda_1^k\left(c_1\mathbf{v}_1 + \sum_{j=2}^{n} c_j\left(\frac{\lambda_j}{\lambda_1}\right)^k\mathbf{v}_j\right),$$

Applying the maximum norm to both sides and using the fact that$||\mathbf{x}_k|| = 1$, we obtain

$$1 = |m_k\lambda_1^k|||\left(c_1\mathbf{v}_1 + \sum_{j=2}^{n} c_j\left(\frac{\lambda_j}{\lambda_1}\right)^k\mathbf{v}_j\right)||_\infty.$$

Consequently, as we take the limit,

$$\lim_{k\to\infty} |m_k\lambda_1^k| = \frac{1}{|c_1|||\mathbf{v}||_\infty} < \infty.$$

This may be expressed as

$$\lim_{k\to\infty} m_k\lambda_1^k = \pm\frac{1}{|c_1|||\mathbf{v}||_\infty} < \infty.$$

Finally,

$$\lim_{k\to\infty} \mathbf{x}_k = \lim_{k\to\infty} m_k \lambda_1^k c_1 \mathbf{v}_1 = K\mathbf{v_1}.$$

Moreover,

$$||\mathbf{x}_k - K_k \mathbf{v}_1||_\infty = \left|\left| \left(m_k \lambda_1^k \sum_{j=2}^{n} c_j \left(\frac{\lambda_j}{\lambda_1}\right)^k \mathbf{v}_j \right) \right|\right|_\infty \leq C \left|\frac{\lambda_2}{\lambda_1}\right|^k$$

For the Rayleigh quotient:

$$\mu_k = \frac{\mathbf{x}_k^T A \mathbf{x}_k}{\mathbf{x}_k^T \mathbf{x}_k}.$$

Since $\mathbf{x}_k \to K\mathbf{v}_1$ and $A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$, we have:

$$\mu_k \to \frac{(K\mathbf{v}_1)^T A (K\mathbf{v}_1)}{(K\mathbf{v}_1)^T (K\mathbf{v}_1)} = \frac{\mathbf{v}_1^T (\lambda_1 \mathbf{v}_1)}{(\mathbf{v}_1^T \mathbf{v}_1)} = \lambda_1.$$

Thus, $\mu_k \to \lambda_1$.