Power Method

Project for Numerical Methods in Calculus and Linear Algebra

Professor: Imre Fekete
Students: Kovács Renáta, Corduneanu-Huci Maria, Dzhekshenova Aiana

GitHub Repository

Why do we need the Power Method?

- The eigenvalues of an $n \times n$ matrix A are obtained by solving its characteristic equation

$$\lambda^n + c_{n-1}\lambda^{n-1} + c_{n-2}\lambda^{n-2} + \cdots + c_0 = 0.$$

- For large n, this polynomial is high degree: difficult to solve and sensitive to rounding.
- Often we only need the eigenvalue with the **largest absolute value**.
- This eigenvalue is called the **dominant eigenvalue**.

Power Method = simple iterative way to approximate this dominant eigenvalue and its eigenvector.

Dominant eigenvalue and eigenvector

Let A have eigenvalues $\lambda_1, \ldots, \lambda_n$.

Dominant eigenvalue

λ_1 is dominant if

$$|\lambda_1| > |\lambda_i| \quad \text{for all } i = 2, \ldots, n.$$

Dominant eigenvector

Any eigenvector corresponding to λ_1 is called a **dominant eigenvector** of A.

Not always!

Some matrices do *not* have a dominant eigenvalue, e.g.

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

Here the largest absolute values are repeated.

Example: finding dominant eigenvalue and eigenvector

Consider

$$A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}.$$

- Characteristic polynomial:

$$\det(A - \lambda I) = (\lambda + 1)(\lambda + 2).$$

- Eigenvalues: $\lambda_1 = -1$, $\lambda_2 = -2$.
- Dominant eigenvalue: $\lambda_{\mathsf{dom}} = -2$ since $|-2| > |-1|$.
- A corresponding eigenvector (solve $(A + 2I)x = 0$):

$$v = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

This is our "target" eigenpair that the Power Method should approximate.

Basic idea of the Power Method

- Assume A has a unique dominant eigenvalue λ_1 with eigenvector e_1.
- Take any non-zero starting vector x_0.
- Form the sequence
$$x_k = Ax_{k-1} = A^k x_0.$$
- When we expand x_0 in the eigenvector basis of A, the term with λ_1^k dominates as k grows.
- So x_k points more and more in direction of e_1.

Repeated multiplication reveals the "strongest direction" of the matrix.

Power Method algorithm

A is a real matrix $n \times n$ with distinct real eigenvalues
Assume:

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$$

$\lambda_1 =$ dominant eigenvalue
$e_1 =$ dominant eigenvector

Method

Given matrix A and starting vector $x_0 \neq 0$:

Repeatedly multiply the initial vector by the matrix:

$$x_0, Ax_0, A^2x_0, A^3x_0, \ldots$$

x_0 can be expressed as a linear combination of the eigenvectors of A:

$$x_0 = c_1 e_1 + c_2 e_2 + \cdots + c_n e_n$$

$$x_1 = Ax_0 = c_1 \lambda_1 e_1 + c_2 \lambda_2 e_2 + \cdots + c_n \lambda_n e_n$$

$$x_p = A^p x_0 = c_1 \lambda_1^p e_1 + c_2 \lambda_2^p e_2 + \cdots + c_n \lambda_n^p e_n$$

Key idea:

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$$

$$\left(\frac{\lambda_i}{\lambda_1}\right)^p$$

shrinks to 0 as $p \to \infty$, for all $i \geq 2$

So

$$x_p \approx c_1 \lambda_1^p e_1$$

and

$$x_{p+1} \approx \lambda_1 x_p,$$

Example: iterations for the 2x2 matrix

Matrix

$$A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}, \quad x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Iteration x_k	Vector x_{k+1}
x_0	$\begin{bmatrix} -10 \\ -4 \end{bmatrix}$
x_1	$\approx \begin{bmatrix} 28 \\ 10 \end{bmatrix}$
$\dots$	$\dots$
x_4	$\approx \begin{bmatrix} -280 \\ -94 \end{bmatrix}$
x_5	$\approx \begin{bmatrix} 568 \\ 190 \end{bmatrix}$

$$x_6 \approx 190 \begin{bmatrix} 2.99 \\ 1 \end{bmatrix}$$

Direction of x_k moves towards eigenvector $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$

Rayleigh quotient: from eigenvector to eigenvalue

The idea

If x is an eigenvector of A with eigenvalue λ, then

$$Ax = \lambda x \quad \Rightarrow \quad x^\top Ax = \lambda x^\top x \quad \Rightarrow \quad \lambda = \frac{x^\top Ax}{x^\top x}.$$

Rayleigh quotient

$$R(x) = \frac{x^\top Ax}{x^\top x}.$$

If x is close to a dominant eigenvector, then $R(x)$ is close to the dominant eigenvalue

So, after running the Power Method, we can plug the last x_k into $R(x_k)$ to approximate $\lambda_{\text{dominant}}$

Example: approximating the eigenvalue

Continue with matrix

$$A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}.$$

Suppose after several iterations we get

$$x \approx \begin{bmatrix} 2.99 \\ 1 \end{bmatrix}$$

$$R(x) = \frac{x^\top A x}{x^\top x}.$$

- Numerically this gives a value very close to -2, which is the dominant eigenvalue.

  Power Method $\Rightarrow$ dominant x; Rayleigh quotient $\Rightarrow$ dominant λ.

Scaling

Without scaling entries of $x_k = A^k x_0$ may grow very large
Two common ways to scale:

1. **Normalize by length:** $x_{k+1} = y_{k+1}/\|y_{k+1}\|$.
2. **Normalize by maximum component:** multiply the result vector by the reciprocal of the largest absolute value inside the vector all components stay between -1 and 1.

Scaling does not change the direction, only the length, so it shows the proportion between the values of the vector.

When does the Power Method converge?

Convergence (informal statement)

Assume:

- A is diagonalizable;
- there is a unique dominant eigenvalue λ_1 with eigenvector v_1;
- starting vector x_0 has a non-zero component in direction v_1.

Then

$$A^k x_0 \ \text{ becomes closer and closer to a multiple of } v_1.$$

So the scaled Power method converges to the dominant eigenvector.

How fast does it converge?

Let eigenvalues be ordered by magnitude:

$$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|.$$

The previous equation for the power method:

$$x_p = A^p x_0 = c_1 \lambda_1^p e_1 + c_2 \lambda_2^p e_2 + \cdots + c_n \lambda_n^p e_n$$

Written a bit differently:

$$x_p = c_1 \lambda_1^p e_1 + \sum_{i=2}^{n} c_i \lambda_i^p e_i = \lambda_1^p \left[c_1 e_1 + \sum_{i=2}^{n} c_i \left(\frac{\lambda_i}{\lambda_1} \right)^p e_i \right].$$

Then the speed of convergence depends on the ratio

$$\rho = \frac{|\lambda_2|}{|\lambda_1|}.$$

- If ρ is small (e.g. 0.1), convergence is **fast**.
- If ρ is close to 1 (e.g. 0.9), convergence is **slow**.

Intuition: terms with λ_2^k die out like ρ^k.

Example: comparing rates

Matrix A

$$A = \begin{bmatrix} -1 & 0 \\ 1 & 6 \end{bmatrix}$$

Eigenvalues: $\lambda_1 = 6$, $\lambda_2 = -1$.
Ratio: $|\lambda_2|/|\lambda_1| = 0.17$
Fast convergence: only few iterations needed.

Matrix B

$$B = \begin{bmatrix} 10 & 0 \\ 0 & 9 \end{bmatrix}$$

Eigenvalues: $\lambda_1 = 10$, $\lambda_2 = 9$.
Ratio: $|\lambda_2|/|\lambda_1| = 0.9$
Slow convergence: many iterations needed.

Same algorithm, very different speed because the eigenvalues are spaced differently.

Real-Life Implementations: PageRank & Beyond

Power Method in Practice

- Used in systems that need to identify the most influential elements within very large networks
- Efficient for very large and sparse matrices because it only requires matrix-vector multiplication

Google PageRank (most famous example)

- Models a "random surfer" on the web.
- Transition matrix represents link structure; damping ($d = 0.15$) allows random jumps.
- PageRank vector = dominant eigenvector of the Google matrix $\widetilde{M}$.
- Computed via the Power Method:

$$v_{k+1} = \widetilde{M} v_k$$

- Scales to billions of webpages due to sparsity.

Other Uses

- Recommendation systems (YouTube, TikTok, Amazon): Personalized PageRank on user–item graphs.

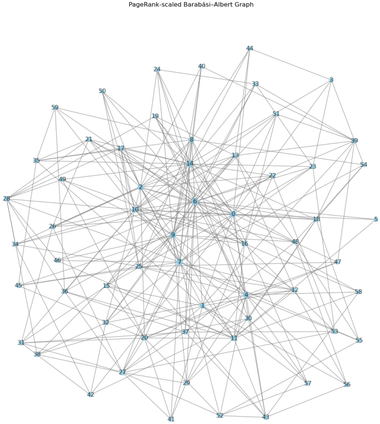

Figure: Google PageRank

Summary

- Dominant eigenvalue: eigenvalue with largest absolute value.
- Power Method:
  - repeatedly multiplies by A,
  - scales vectors to control size,
  - converges to the dominant eigenvector.
- Rayleigh quotient turns an approximate eigenvector into an approximate eigenvalue.
- Convergence is guaranteed under reasonable assumptions and is faster when $|\lambda_2|/|\lambda_1|$ is small.

Takeaway: simple algorithm, powerful for large problems when one eigenpair is enough.