The PageRank Algorithm

Edie Miglio

What is PageRank?

PageRank is an algorithm used by Google Search to rank web pages in their search engine results. It's a way of measuring the "importance" of website pages.

The "Random Surfer" Model

Imagine a user randomly clicking on links.

- The user starts on a random page.
- At each step, the user randomly clicks on one of the links on the current page and moves to the next page.

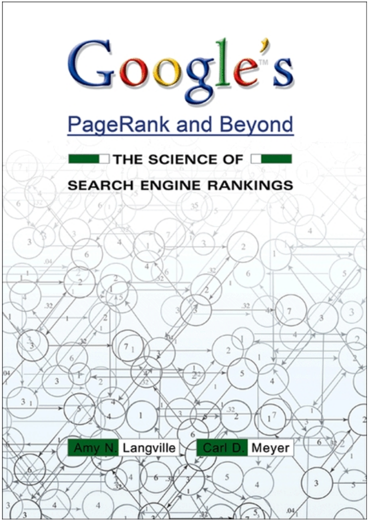
A page is considered more important if a random surfer is more likely to land on it. The PageRank of a page is its long-term visit probability.

Key Insight

A link from an important page is a more significant "vote" than a link from a minor page.

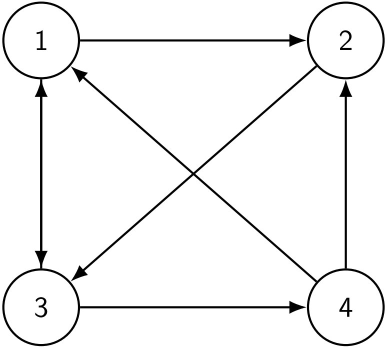
Page Rank

In this presentation we are going to introduce a simplified version of the algorithm. For a detailed explanation see:



A Simple Network Example

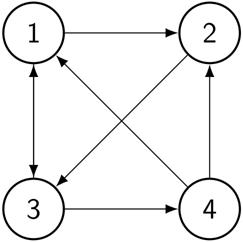
Let's model a small web with 4 pages. The arrows represent hyperlinks.



We will use this network to build the mathematical model.

The Adjacency Matrix A

We can represent the link structure with an **adjacency matrix** A, where $A_{ij} = 1$ if there is a link from page j to page i.



$$\mathrm{A} = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

For example, $A_{21} = 1$ because there is a link from page 1 to 2.

The Transition Matrix M

The adjacency matrix shows connections, but we need probabilities. The **transition matrix** M contains the probability of moving from page j to page i.

Construction

We normalize each column of A by dividing it by its sum (the number of outgoing links from that page).

For our example:

- Page 1 has 2 outgoing links.
- Page 2 has 1 outgoing link.
- Page 3 has 2 outgoing links.
- Page 4 has 2 outgoing links.

$$\mathsf{M} = \begin{pmatrix} 0 & 0 & 1/2 & 1/2 \\ 1/2 & 0 & 0 & 1/2 \\ 1/2 & 1 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 \end{pmatrix}$$

M is a **column stochastic matrix** (all columns sum to 1).

The State Vector π

Let $\pi_k \in \mathbb{R}^4$ be a vector representing the probability distribution of our random surfer at step k. The i-th component, $(\pi_k)_i$, is the probability that the surfer is on page i.

The state evolves according to the rule:

$$\pi_{k+1} = \mathsf{M}\pi_k$$

The Steady State

As $k \to \infty$, this process converges to a stationary probability distribution π, where the probabilities no longer change. This steady state vector must satisfy:

$$\pi = \mathsf{M}\pi$$

This vector π is the **PageRank vector**.

Connection to Eigenvectors

The steady-state equation is an eigenvector problem!

$$\mathsf{M}\pi = \pi \quad \Longleftrightarrow \quad \mathsf{M}\pi = 1 \cdot \pi$$

The PageRank Vector

The PageRank vector π is the eigenvector of the transition matrix M corresponding to the eigenvalue $\lambda = 1$.

Key Questions

1. Is $\lambda = 1$ always an eigenvalue of M?
2. Is this eigenvalue the largest one?
3. Is the corresponding eigenvector unique?

Guarantees: Perron-Frobenius Theorem

The Perron-Frobenius theorem for stochastic matrices gives us the answers.

Theorem (Perron-Frobenius, simplified)

If M *is a column stochastic, positive, and irreducible matrix* [a]*, then:*

1. *The eigenvalue* $\lambda = 1$ *is the largest eigenvalue (the "principal eigenvalue").*
2. *The eigenvector corresponding to* $\lambda = 1$ *is unique (up to a scalar multiple).*
3. *This eigenvector can be chosen to have strictly positive components.*

[a] An irreducible stochastic matrix is one where every state can be reached from every other state in a finite number of steps. For a stochastic matrix P to be irreducible, there must exist an n such that $P^n(i,j) > 0$ for all states i and j.

In Practice

The "Google Matrix" (a slightly modified version of M) is constructed to satisfy these properties, ensuring a unique and positive PageRank vector exists.

The Power Method Algorithm

Since we know that the PageRank vector is the eigenvector associated with the largest eigenvalue ($\lambda = 1$), we can compute it efficiently using the **Power Method**.

The Algorithm

1. Start with an initial guess for the probability vector, π_0. A uniform distribution is a common choice:

$$\pi_0 = [1/N, 1/N, \ldots, 1/N]^T$$

2. Iterate until convergence:

$$\pi_{k+1} = \mathrm{M}\pi_k$$

3. Stop when the change between π_{k+1} and π_k is very small.

This is simple, scalable, and guaranteed to converge to the PageRank vector.

Why the Power Method Converges

Let the eigenvectors of M be $\mathrm{v}_1, \mathrm{v}_2, \dots, \mathrm{v}_N$ with corresponding eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_N$. From Perron-Frobenius, we know $\lambda_1 = 1$ is the largest: $1 = |\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_N|$.

1. **Decomposition:** We can write our initial guess π_0 as a linear combination of the eigenvectors:

$$\pi_0 = c_1\mathrm{v}_1 + c_2\mathrm{v}_2 + \dots + c_N\mathrm{v}_N$$

2. **Iteration:** After one step, we have:

$$\pi_1 = \mathrm{M}\pi_0 = c_1\lambda_1\mathrm{v}_1 + c_2\lambda_2\mathrm{v}_2 + \dots$$

   After k steps:

$$\pi_k = \mathrm{M}^k\pi_0 = c_1\lambda_1^k\mathrm{v}_1 + c_2\lambda_2^k\mathrm{v}_2 + \dots$$

3. **Convergence:** Since $\lambda_1 = 1$ and $|\lambda_i| < 1$ for $i > 1$, as $k \to \infty$, the terms λ_i^k go to zero much faster than λ_1^k.

$$\pi_k \approx c_1(1)^k\mathrm{v}_1 = c_1\mathrm{v}_1$$

   The vector converges to a multiple of the dominant eigenvector v_1, which is our PageRank vector π.

Appendix: Proof for Perron-Frobenius (Simplified)

We want to prove two key properties for a column stochastic matrix M.

1. $\lambda = 1$ is an eigenvalue of M.
2. For any eigenvalue λ of M, its magnitude is bounded: $|\lambda| \leq 1$.

Together, these points imply that $\lambda = 1$ is the largest (or dominant) eigenvalue.

Proof Part 1: $\lambda = 1$ is an eigenvalue I

By definition, $\lambda = 1$ is an eigenvalue if the matrix $(\mathrm{M} - 1 \cdot \mathrm{I})$ is singular, i.e., if its determinant is zero.

Key Property

A matrix is singular if its columns (or rows) are linearly dependent.

1. Since M is column stochastic, the sum of the entries in each column is 1.
$$\sum_{i=1}^{N} M_{ij} = 1 \quad \forall j$$
2. Let $1^T = [1, 1, \ldots, 1]$. Then $1^T \mathrm{M} = 1^T$.
3. Rearranging this gives: $1^T \mathrm{M} - 1^T = 0^T \implies 1^T(\mathrm{M} - \mathrm{I}) = 0^T$.
4. This equation shows that the sum of the rows of $(\mathrm{M} - \mathrm{I})$ is the zero vector. Therefore, the rows are linearly dependent.

Proof Part 1: $\lambda = 1$ is an eigenvalue II

5. Since the rows are linearly dependent, the matrix $(\mathsf{M} - \mathsf{I})$ is singular, and $\det(\mathsf{M} - \mathsf{I}) = 0$.

Conclusion

This proves that $\lambda = 1$ is always an eigenvalue of M.

Proof Part 2: $|\lambda| \leq 1$ for all eigenvalues I

We use the property that for any eigenvalue λ, its magnitude is bounded by any induced matrix norm: $|\lambda| \leq \|\mathrm{M}\|$.

The 1-Norm

The induced matrix 1-norm is defined as the maximum absolute column sum:

$$\|\mathrm{M}\|_1 = \max_j \sum_{i=1}^{N} |M_{ij}|$$

1. For a column stochastic matrix M, all entries are non-negative ($M_{ij} \geq 0$), so $|M_{ij}| = M_{ij}$.
2. The sum of each column is exactly 1.

$$\sum_{i=1}^{N} M_{ij} = 1 \quad \forall j$$

Proof Part 2: $|\lambda| \leq 1$ for all eigenvalues II

3. Therefore, the maximum column sum is also 1.

$$\|M\|_1 = \max_j(1) = 1$$

Conclusion

Since $|\lambda| \leq \|M\|_1$, we have $|\lambda| \leq 1$ for any eigenvalue λ of M. This proves that no eigenvalue can have a magnitude greater than 1.