Lecture 11: The PageRank Algorithm

Modelling Complex Systems (Spring 2026)

Jordi-Lluís Figueras

Matematiska Institutionen, Uppsala Universitet

April 28, 2026

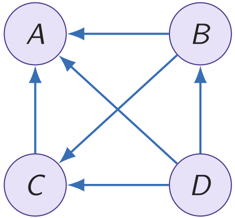
Prelecture discussion

Before Lecture 11, think about the web as a directed network of pages.

- A page can be important because many pages link to it.
- But a link from an important page should probably count more than a link from an obscure page.
- If a page links to many places, each individual outgoing link should probably receive only part of its influence.

Questions to reflect on:

- Can importance be defined recursively?
- Why is this naturally a matrix problem?
- What can go wrong if a page has no outgoing links?



links between four web pages

Lecture goals

In this lecture we will:

- recap Markov chains as linear dynamics on directed weighted graphs;
- connect stationary distributions with eigenvectors for eigenvalue 1;
- state the Perron-Frobenius message needed for convergence;
- motivate PageRank as an importance score on a directed web graph;
- formulate the simplified PageRank rule;
- compute the first iteration for the four-page example A, B, C, D;
- rewrite the rule as a matrix iteration;
- identify why dangling pages and reducible link structures force refinements.

Markov-Chain Recap

Where we stopped

In Lecture 10 we introduced networks and Markov chains on directed weighted graphs.

The city example gave a transition matrix P and a population vector x_t.

The update rule was

$$x_{t+1} = Px_t.$$

The simulation showed this repeated redistribution of population across the network. Today we first complete the linear-algebra interpretation of that behavior.

Transition matrices

For the convention used in these lectures, P_{ij} is the probability of moving from node j to node i in one step.

Thus P is column-stochastic:

$$P_{ij} \geq 0, \qquad \sum_{i=1}^{n} P_{ij} = 1 \quad \text{for each } j.$$

Each column describes how the mass currently located at one node is distributed among possible destinations.

Zeros in P encode forbidden transitions in the directed graph.

Mass preservation

Because each column of P sums to 1, total mass is preserved:

$$\sum_{i=1}^{n}(x_{t+1})_i = \sum_{i=1}^{n}\sum_{j=1}^{n} P_{ij}(x_t)_j = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} P_{ij}\right)(x_t)_j = \sum_{j=1}^{n}(x_t)_j.$$

If the total mass is normalized to 1, then x_t is a probability distribution on the nodes.

The Markov chain is therefore a linear dynamical system on the probability simplex.

Stationary distributions

An equilibrium for the update $x_{t+1} = Px_t$ is a vector x^* satisfying

$$Px^* = x^*.$$

So a stationary state is exactly an eigenvector with eigenvalue 1.

If x^* has nonnegative entries and

$$\sum_i x_i^* = 1,$$

then x^* is called a stationary distribution.

This is the bridge from stochastic dynamics to spectral theory.

Perron-Frobenius

Why eigenvalues matter

Very roughly, powers of a matrix are controlled by its eigenvalues.

If all eigenvalues other than 1 have modulus strictly smaller than 1, then repeated multiplication by P damps out those components.

So one expects

$$P^k x_0 \longrightarrow x^* \qquad \text{as } k \to \infty,$$

where x^* is a stationary distribution.

This is the basic mechanism behind convergence to equilibrium in many network-transport models.

Possible obstructions

Convergence is not automatic for every stochastic matrix. Typical obstructions are:

- **reducibility:** the graph splits into disconnected communicating parts;
- **periodicity:** the dynamics cycles between classes of states;
- **multiple closed classes:** different long-time outcomes depend on the initial condition.

So we need structural assumptions on the graph and positivity assumptions on the matrix. Those assumptions are precisely where Perron-Frobenius enters.

Perron-Frobenius theorem: informal message

The Perron-Frobenius theorem applies to matrices with nonnegative entries. It tells us that positivity forces a strong spectral structure.

For Markov chains, the practical message is:

- there is a distinguished dominant eigenvalue;
- it has a nonnegative eigenvector;
- under stronger positivity assumptions, that dominant direction is unique and attracts the dynamics.

This is exactly what we need to justify convergence to a unique long-time population profile.

Perron-Frobenius theorem

Perron-Frobenius theorem for a positive matrix

Let $A \in \mathbb{R}^{n \times n}$ have strictly positive entries: $A_{ij} > 0$ for all i, j. Then:

1. A has a real eigenvalue $\rho(A) > 0$ equal to its spectral radius;
2. there is an eigenvector v associated with $\rho(A)$ whose entries are all strictly positive;
3. the eigenvalue $\rho(A)$ is simple;
4. every other eigenvalue λ satisfies $|\lambda| < \rho(A)$.

There are more general versions for irreducible or primitive nonnegative matrices, which are the forms most relevant to Markov chains.

What the theorem says for stochastic matrices

For a stochastic matrix P, the special eigenvalue is already known:

$$\rho(P) = 1.$$

Indeed, total mass is preserved, so 1 is the natural scale of the dynamics.

If P is sufficiently connected and aperiodic, Perron-Frobenius implies:

- there exists a unique stationary distribution x^*;
- x^* has strictly positive entries;
- for every initial distribution x_0,

$$P^k x_0 \to x^*.$$

So the dynamics forgets the initial condition and settles to a network-dependent equilibrium.

Back to the city interpretation

In the city example, a stationary distribution means:

after one more day of movement, the expected population in each neighborhood is unchanged.

This does *not* mean that individuals stop moving. It means that the aggregate inflow and outflow balance at each node.

That is a recurring theme in complex systems:

Microscopic motion, macroscopic stationarity

Agents may keep moving all the time, while the large-scale distribution becomes stable.

Ranking Pages

Bridge from Lecture 10

We have now completed the missing bridge from Lecture 10:

Markov chains on networks lead to stationary distributions, and Perron-Frobenius explains when these stationary distributions are unique and attracting.

The PageRank algorithm is one of the most famous applications of this idea.

The modelling move is:

Network interpretation

Pages are nodes, hyperlinks are directed edges, and page importance is interpreted through the long-time behavior of movement on this directed graph.

Thus PageRank is not only a search-engine idea. It is also a clean example of spectral ranking on a complex network.

The basic problem

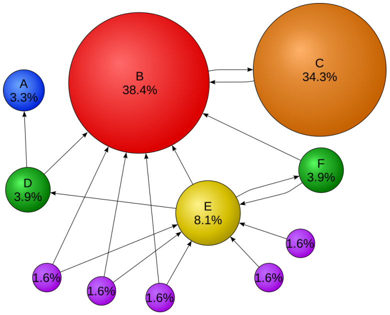
Suppose we have many web pages and hyperlinks between them.

We want to assign each page u a number $PR(u)$ measuring its importance.

Simple link counting is not enough:

- a link from an important page should carry more weight;
- a page that links to many pages should split its influence;
- importance should be global, not only local.

This leads to a recursive definition: a page is important if important pages link to it.



Source: Wikimedia Commons, public domain.

Votes, but weighted votes

PageRank can be introduced through a voting metaphor.

Each page distributes its current rank equally among the pages it links to.

If page v has $L(v)$ outgoing links, then each outgoing link from v contributes

$$\frac{PR(v)}{L(v)}.$$

Therefore the rank of a page is obtained by summing the contributions from all pages that point to it.

Simplified PageRank rule

Let B_u be the set of pages that link to page u. Let $L(v)$ be the number of outgoing links from page v.

The simplified PageRank rule is

$$PR(u) = \sum_{v \in B_u} \frac{PR(v)}{L(v)}.$$

This is a recursive equation, because the unknown ranks appear on both sides.

Computationally, we first use it as an iteration:

$$PR_{t+1}(u) = \sum_{v \in B_u} \frac{PR_t(v)}{L(v)}.$$

Initial condition

For a small network with N pages, a natural initial condition is the uniform distribution

$$PR_0(u) = \frac{1}{N}.$$

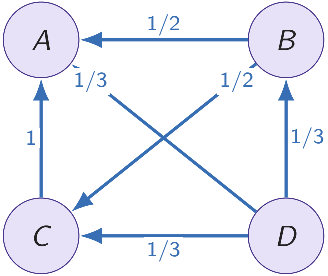
For the four-page example A, B, C, D,

$$PR_0(A) = PR_0(B) = PR_0(C) = PR_0(D) = \frac{1}{4}.$$

The iteration then repeatedly redistributes rank along outgoing links.

The Four-Page Example

A visual graph of A, B, C, D



The links are:

- B links to A and C;
- C links to A;
- D links to A, B, and C;
- A has no outgoing link in this simplified example.

The edge labels show how the current rank is split among outgoing links.

Computing the first update

Start from the uniform values

$$PR_0(A) = PR_0(B) = PR_0(C) = PR_0(D) = \frac{1}{4}.$$

For page A, the incoming pages are B, C, D. Therefore

$$\begin{aligned} PR_1(A) &= \frac{PR_0(B)}{2} + \frac{PR_0(C)}{1} + \frac{PR_0(D)}{3} \\ &= \frac{1/4}{2} + \frac{1/4}{1} + \frac{1/4}{3} = \frac{11}{24} \approx 0.458. \end{aligned}$$

The rank of A becomes large because it receives support from three pages, including all of C's rank.

The other pages after one update

Using the same rule:

$$PR_1(B) = \frac{PR_0(D)}{3} = \frac{1}{12},$$

$$PR_1(C) = \frac{PR_0(B)}{2} + \frac{PR_0(D)}{3} = \frac{1}{8} + \frac{1}{12} = \frac{5}{24},$$

$$PR_1(D) = 0.$$

So the first updated vector, in the order (A, B, C, D), is

$$r_1 = \begin{pmatrix} 11/24 \\ 1/12 \\ 5/24 \\ 0 \end{pmatrix}.$$

A warning: mass has disappeared

In the first update we obtained

$$\frac{11}{24} + \frac{1}{12} + \frac{5}{24} + 0 = \frac{3}{4}.$$

The missing mass is not a calculation error. It comes from page A.

Since A has no outgoing links, its initial rank $1/4$ was not passed on to any page.

Such a page is called a dangling page or sink. This is one reason the practical PageRank algorithm needs an additional correction.

Matrix Formulation

The link matrix

Let r_t be the column vector of PageRank values in the order (A, B, C, D).

The simplified update is

$$r_{t+1} = Mr_t,$$

where

$$M = \begin{pmatrix} 0 & 1/2 & 1 & 1/3 \\ 0 & 0 & 0 & 1/3 \\ 0 & 1/2 & 0 & 1/3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The entry M_{ij} is the fraction of page j's rank that flows to page i.

Column interpretation

Each column describes how one page distributes its rank.

For example:

- column B has entries $1/2$ in rows A and C;
- column C has entry 1 in row A;
- column D has entries $1/3$ in rows A, B, and C;
- column A is zero because A has no outgoing links.

If every page had at least one outgoing link and all rank were distributed, then M would be column-stochastic.

Connection with Markov chains

If M is column-stochastic, then the update

$$r_{t+1} = Mr_t$$

is exactly the Markov-chain update from Lecture 10.

The random-surfer interpretation is:

- a surfer currently on page j chooses one of its outgoing links uniformly;
- the probability distribution over pages evolves by multiplying by M;
- the long-time distribution gives the PageRank scores.

Thus PageRank is a stationary distribution problem on a directed web graph.

The eigenvector viewpoint

At convergence, a PageRank vector r^* should satisfy

$$Mr^* = r^*.$$

So PageRank is an eigenvector centrality:

Key equation

The PageRank vector is a nonnegative eigenvector associated with eigenvalue 1 for a suitable transition matrix.

This is why the Perron-Frobenius theorem from Lecture 10 is directly relevant. It explains when a positive, unique, attracting rank vector exists.

What the simplified model misses

The simplified rule is pedagogically useful, but it is not yet the robust PageRank algorithm.

Two problems appear immediately:

- **dangling pages:** pages with no outgoing links trap or remove mass;
- **disconnected components:** the web graph may not be strongly connected.

The standard correction is to add a small probability of jumping to a random page instead of following a link. This is the damping factor or random teleportation idea.

Preview: damping and teleportation

In the damped PageRank model, the surfer does two things:

- with probability d, follow an outgoing link;
- with probability $1 - d$, jump to a uniformly chosen page.

The usual value is around $d = 0.85$.

In matrix form, this will become

$$r_{t+1} = dMr_t + \frac{1-d}{N}\mathbf{1}.$$

We will use this refinement to turn the simplified idea into a well-behaved Markov chain.

What to remember

- PageRank assigns importance to nodes in a directed link network.
- The simplified rule says that a page receives rank from pages linking to it, split by their number of outgoing links.
- The four-page example A, B, C, D shows explicitly how rank is redistributed.
- In matrix form, the update is $r_{t+1} = Mr_t$.
- PageRank is fundamentally a stationary-distribution or eigenvector problem.
- Dangling pages and poor connectivity motivate damping and teleportation.

Suggested bibliography

For students who want to read more:

- S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Computer Networks and ISDN Systems*, 30, 107–117, 1998.
- L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank citation ranking: bringing order to the web," Stanford InfoLab technical report, 1999.
- A. N. Langville and C. D. Meyer, *Google's PageRank and Beyond: The Science of Search Engine Rankings*, Princeton University Press, 2006.
- Wikipedia contributors, "PageRank," especially the simplified algorithm section.

Acknowledgment

These slides were prepared with the assistance of OpenAI Codex / ChatGPT 5.5 as a drafting and editing tool.
All mathematical, pedagogical, and course-specific content remains under the responsibility of the lecturer.