A linear map of the web

Probability (and techniques)

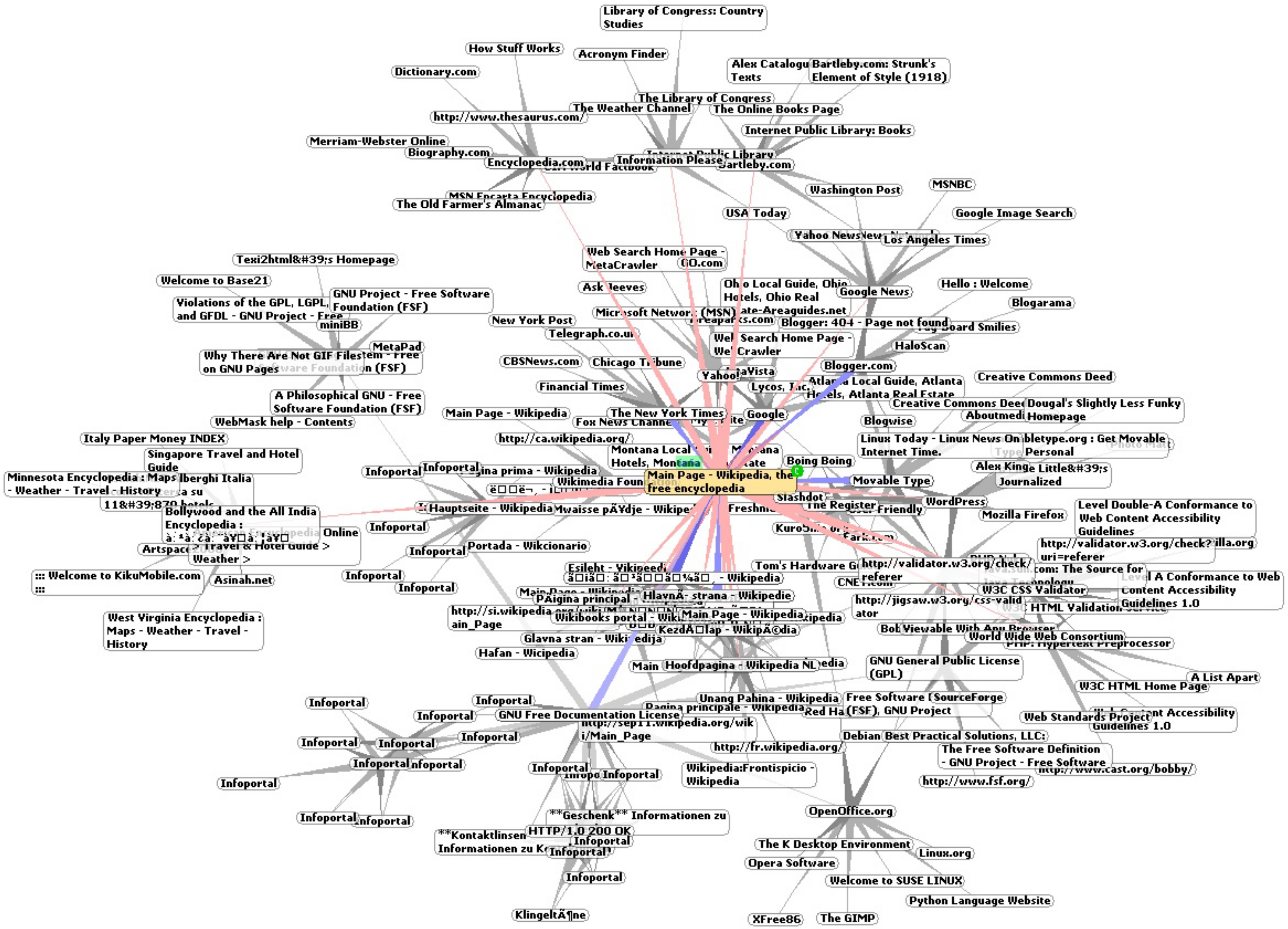
Random Walks

Stationary Distribution

Perron-Frobenius Theorem

Algorithms (models and applications)

Pagerank



Given search query:

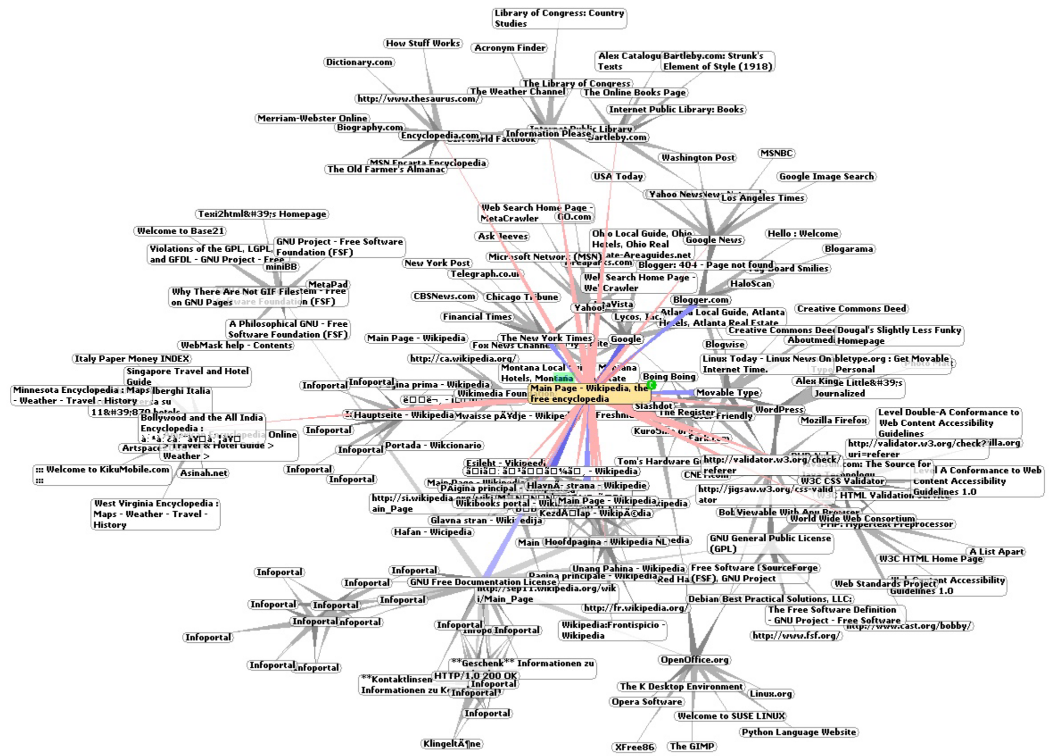
① what are the relevant webpages?

② what are the most relevant webpages?

② most relevant webpages?

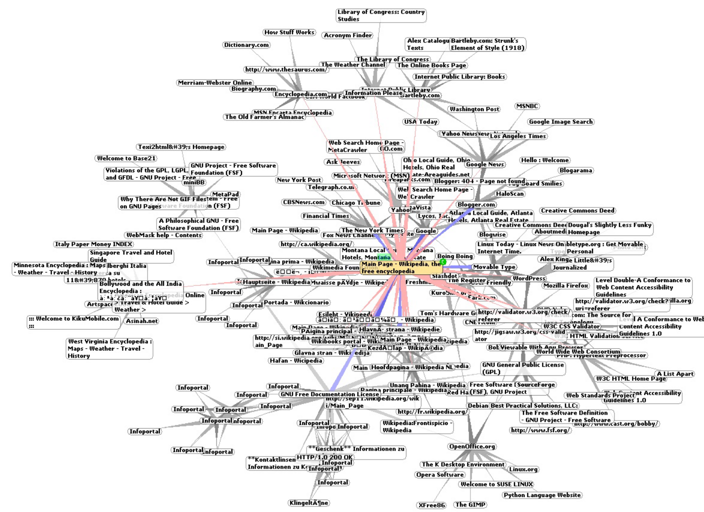
this is a ranking question

ideas?



Idea 1:

$$\mathrm{score}_1(v) = \sum_{(u,v)\in E} 1$$



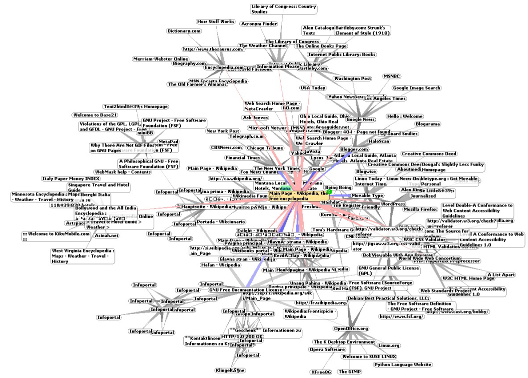
Idea 2:

$$\mathrm{score}_2(v) = \sum_{(u,v)\in E} \frac{1}{\deg^+(u)}$$

$\deg^+(u)$ = out-degree of u

Idea 3:

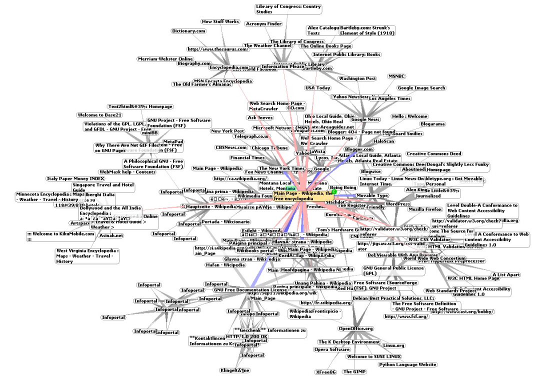
$$\mathrm{score}_3(v) = \sum_{(u,v)\in E} \frac{\mathrm{score}_3(u)}{\deg^+(u)}$$



$$\boxed{\mathrm{score}_3(v) = \sum_{(u,v)\in E} \frac{\mathrm{score}_3(u)}{\deg^+(u)}}$$

Not clear that $\exists$ values $\mathrm{score}_3(v)$ satisfying $\forall v$

$$\boxed{\mathrm{score}_3(v) = \sum_{(u,v)\in E} \frac{\mathrm{score}_3(u)}{\deg^+(u)}}$$



let $\Delta^V = \{x \in \mathbb{R}_V^{\geq 0} : \sum_v x_v = 1\} = \{\text{distributions over } V\}$

Theorem there exists $x \in \Delta^V$ s.t.

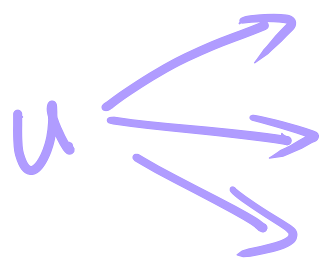
$$x_v = \sum_{(u,v)\in E} \frac{x_u}{d^+(u)} \quad \forall v \in V.$$

If G is strongly connected, then x is unique and strictly positive.

(why?)

Random walks

from a vertex u, pick an edge (u,v) leaving u randomly, and go to v. (repeat)

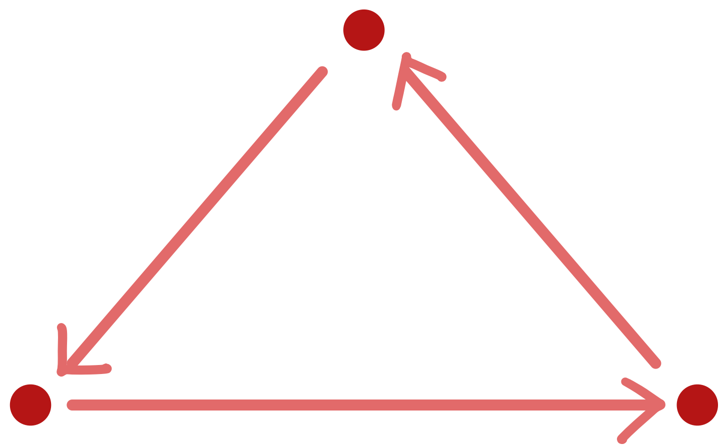


After k steps, we have different probabilities $p_v^{(k)}$ of being at various vertices v

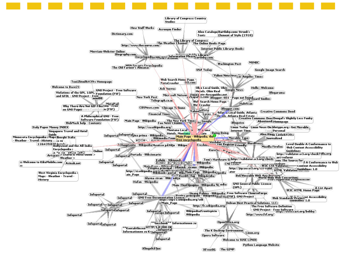
Q: do the $p_v^{(k)}$'s converge as $k \to \infty$?

Stationary distribution

$x \in \Delta^V$ is a stationary distribution if taking a random step from x gives x



$$\boxed{\text{score}_3(v) = \sum_{(u,v)\in E} \frac{\text{score}_3(u)}{\deg^+(u)}}$$



let $\Delta^V = \{x \in \mathbb{R}_V^{\geq 0} : \sum_v x_v = 1\} = \{\text{distributions over } V\}$

Theorem there exists $x \in \Delta^V$ s.t.

$$x_v = \sum_{(u,v)\in E} \frac{x_u}{d^+(u)} \quad \forall\, v \in V.$$

If G is strongly connected, then x is unique and strictly positive.

Theorem, restated: every random walk has a stationary distribution. If G is strongly connected, then this distribution is unique.

why? linear algebra!

Linear algebra basics

$f:\mathbb{R}^n \to \mathbb{R}^n$ is linear if

$$f(x+y) = f(x) + f(y)$$

$$f(\alpha x) = \alpha f(x)$$

All linear $f:\mathbb{R}^n \to \mathbb{R}^n$ has the form

$$(f(x))_i = (Ax)_i = \sum_j A_{ij} x_j \quad \text{for unique } A \in \mathbb{R}^{n \times n}$$

Adjacency matrix / linear map

$A : \mathbb{R}^V \to \mathbb{R}^V$ where

$$\langle e_v, A e_u \rangle = A_{vu} = \begin{cases} 1 & \text{if } (u,v) \in E \\ 0 & \text{otherwise} \end{cases}$$

(e_u = $\{0,1\}$-indicator vector for $u \in V$)

e.g. $\mathrm{score}_1(v) = \sum_{(u,v) \in E} 1 =$

$\deg^+(u) =$

Random walk matrix / linear map

$$\langle e_v, Ae_u \rangle = A_{vu} = \begin{cases} 1 & \text{if } (u,v) \in E \\ 0 & \text{otherwise} \end{cases}$$

e.g. $\text{score}_1(v) = \sum_{(u,v)\in E} 1 = \langle e_v, A\mathbb{1} \rangle$

let $D = \text{diag}(\deg^+)$

Random walk matrix is $R = AD^{-1}$

$$\langle e_v, Re_u \rangle = R_{vu} = \sum_w A_{vw} D^{-1}_{wu} = \frac{A_{vu}}{\deg^+(u)}$$

e.g. $\text{score}_2(v) = \langle e_v, R\mathbb{1} \rangle$

stationary dist. x satisfies $x = Rx$

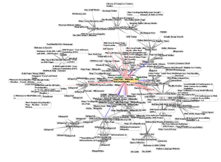
$\boxed{\text{stationary dist. } x \text{ satisfies } x = Rx}$

"eigenvalue"

$$Ax = \lambda x$$

"eigenvector"

$$\mathrm{score}_3(v) = \sum_{(u,v)\in E} \frac{\mathrm{score}_3(u)}{\deg^+(u)}$$



let $\Delta^V = \{x \in \mathbb{R}_V^{\geq 0} : \sum_v x_v = 1\} = \{\text{distributions over } V\}$

Theorem there exists $x \in \Delta^V$ s.t.

$$x_v = \sum_{(u,v)\in E} \frac{x_u}{d^+(u)} \quad \forall\, v \in V.$$

If G is strongly connected, then x is unique and strictly positive.

Theorem, restated: every random walk has a stationary distribution. If G is strongly connected, then this distribution is unique.

Theorem, restated: let G be a dir. graph and R the random walk linear map. Then R has an eigenvector $x \in \Delta^V$ w/ eigenvalue 1. If G is strongly connected, then x is the unique eigenvector w/ eigenvalue 1

$\boxed{\text{stationary dist. } x \text{ satisfies } x = Rx}$

"eigenvalue"

$$Ax = \lambda x$$

"eigenvector"

stationary dist. x satisfies $x = Rx$

$$A x = \lambda x$$

"eigenvalue" "eigenvector"

Lemma every linear map A has an eigenvalue $\lambda \in \mathbb{C}$ and eigenvector $x \in \mathbb{C}^n$

Proof fix v nonzero

v, Av, $A^2 v$, $A^3 v$, $\vdots$, $A^{n-1} v$, $A^n v$

$n+1$ vectors $\Rightarrow$ linear dependent

$\Rightarrow \alpha_0 v + \alpha_1 A v + \cdots + \alpha_n A^n v = 0$ for α not all zero

i.e. $p(A)v = 0$

for polynomial $p(x)$, $\deg(p) = n$

$$p(x) = (x - r_1)(x - r_2)\cdots(x - r_n) \qquad r_1, \ldots, r_n \in \mathbb{C}$$

$$p(A)v = (A - r_1 I)(A - r_2 I)\cdots(A - r_n) I = v$$

Fundamental Theorem of Algebra
polynomial $p(x)$ has $\deg(p)$ roots

Linear algebra basics

$f : \mathbb{R}^n \to \mathbb{R}^n$ is linear if

$$f(x+y) = f(x) + f(y)$$

$$f(\alpha x) = \alpha f(x)$$

All linear $f : \mathbb{R}^n \to \mathbb{R}^n$ has the form

$$(f(x))_i = (Ax)_i = \sum_j A_{ij} x_j \quad \text{for unique } A \in \mathbb{R}^{n \times n}$$

Transpose for $A \in \mathbb{R}^{n \times n}$

$$\langle x, Ay \rangle = \langle A^T x, y \rangle \quad \forall x, y$$

Let $A: \mathbb{R}^n \to \mathbb{R}^n$ linear

stationary dist. x satisfies $x = Rx$

Lemma:

A invertible $\Leftrightarrow$ A^T invertible

"eigenvalue"

$$Ax = \lambda x$$

"eigenvector"

Lemma every linear map A has an eigenvalue $\lambda \in \mathbb{C}$ and eigenvector $x \in \mathbb{C}^n$

Let $A: \mathbb{R}^n \to \mathbb{R}^n$ linear

$\boxed{\text{stationary dist. } x \text{ satisfies } x = Rx}$

Lemma: A and A^T have same eigenvalues w/ same multiplicities

"eigenvalue"

$$Ax = \lambda x$$

"eigenvector"

Lemma every linear map A has an eigenvalue $\lambda \in \mathbb{C}$ and eigenvector $x \in \mathbb{C}^n$

Lemma:
A invertible $\iff$ A^T invertible

Let $A: \mathbb{R}^n \to \mathbb{R}^n$ linear

$\boxed{\text{stationary dist. } x \text{ satisfies } x = Rx}$

A is "**positive**" if

$\langle x, Ay \rangle > 0 \quad \forall$ nonzero $x, y \in \mathbb{R}^n_{\geq 0}$

"eigenvalue"

$$Ax = \lambda x$$

"eigenvector"

Lemma every linear map A has an eigenvalue $\lambda \in \mathbb{C}$ and eigenvector $x \in \mathbb{C}^n$

Lemma: A invertible $\Leftrightarrow$ A^T invertible

Lemma: A and A^T have same eigenvalues w/ same multiplicities

$\boxed{\text{Perron-Frobenius Theorem}}$: If A is positive, then A has eigenvalue λ_1, eigenvector x_1 s.t.

(a) $\lambda_1 > 0$ and $x_1 > 0$

(b) x_1 is unique eigenvector of λ_1

(c) x_1 is unique nonnegative vector s.t. $Ax_1 \geq \lambda_1 x_1$

(d) All other eigenvalues μ have $|\mu| < \lambda_1$

(e) No other eigenvectors are nonnegative

Proof let

stationary dist. x satisfies $x = Rx$

$$L = \{\lambda > 0 \text{ s.t. } Ax \geq \lambda x \text{ for some } x \in \Delta^n\}$$

$$\underbrace{A}_{} \overbrace{}^{\text{"eigenvalue"}} x = \lambda x$$

"eigenvalue"

$Ax = \lambda x$

"eigenvector"

approach:

supremum λ^+ of $L = \lambda_1$

Lemma every linear map A has an eigenvalue $\lambda \in \mathbb{C}$ and eigenvector $x \in \mathbb{C}^n$

Lemma:
A invertible $\Leftrightarrow$ A^T invertible

Lemma: A and A^T have same eigenvalues w/ same multiplicities

A is "positive" if $\langle x, Ay \rangle > 0$ $\forall$ nonzero $x, y \in \mathbb{R}^n_{\geq 0}$

Theorem: If A is positive, then A has eigenvalue λ_1, eigenvector x_1 s.t.

(a) $\lambda_1 > 0$ and $x_1 > 0$

(b) x_1 is unique eigenvector of λ_1

(c) x_1 is unique nonnegative vector s.t. $Ax_1 \geq \lambda_1 x_1$

(d) All other eigenvalues μ have $|\mu| < \lambda_1$

(e) No other eigenvectors are nonnegative

stationary dist. x satisfies $x = Rx$

"eigenvalue"

$$A x = \lambda x$$

"eigenvector"

Lemma every linear map A has an eigenvalue $\lambda \in \mathbb{C}$ and eigenvector $x \in \mathbb{C}^n$

Proof fix v nonzero

v

Av

$A^2 v$

$A^3 v$

$\vdots$

$A^{n-1} v$

$A^n v$

$n+1$ vectors $\Rightarrow$ linear dependent

$\Rightarrow \alpha_0 v + \alpha_1 A v + \cdots + \alpha_n A^n v = 0$ for α not all zero

i.e. $p(A) v = 0$

for polynomial $p(x)$, $\deg(p) = n$

$$p(x) = (x - r_1)(x - r_2) \cdots (x - r_n) \qquad r_1, \ldots, r_n \in \mathbb{C}$$

$$p(A) v = (A - r_1 I)(A - r_2 I) \cdots (A - r_n) I = v$$

Fundamental Theorem of Algebra

polynomial $p(x)$ has $\deg(p)$ roots

Linear algebra basics

$f:\mathbb{R}^n \to \mathbb{R}^n$ is linear if

$$f(x+y) = f(x) + f(y)$$

$$f(\alpha x) = \alpha f(x)$$

All linear $f:\mathbb{R}^n \to \mathbb{R}^n$ has the form

$$(f(x))_i = (Ax)_i = \sum_j A_{ij} x_j \quad \text{for unique } A \in \mathbb{R}^{n \times n}$$

Transpose for $A \in \mathbb{R}^{n \times n}$

$$\langle x, Ay \rangle = \langle A^T x, y \rangle \quad \forall x, y$$

Proof let

stationary dist. x satisfies $x = Rx$

$$L = \{\lambda > 0 \text{ s.t. } Ax \geq \lambda x \text{ for some } x \in \Delta^n\}$$

"eigenvalue"

$$Ax = \lambda x$$

"eigenvector"

Lemma every linear map A has an eigenvalue $\lambda \in \mathbb{C}$ and eigenvector $x \in \mathbb{C}^n$

Lemma: A invertible $\Leftrightarrow$ A^T invertible

Lemma: A and A^T have same eigenvalues w/ same multiplicities

A is "**positive**" if $\langle x, Ay \rangle > 0 \ \forall$ nonzero $x, y \in \mathbb{R}^n_{\geq 0}$

Theorem: If A is positive, then A has eigenvalue λ_1, eigenvector x_1 s.t.

(a) $\lambda_1 > 0$ and $x_1 > 0$

(b) x_1 is unique eigenvector of λ_1

(c) x_1 is unique nonnegative vector s.t. $Ax_1 \geq \lambda_1 x_1$

(d) All other eigenvalues μ have $|\mu| < \lambda_1$

(e) No other eigenvectors are nonnegative

claims:

- L is nonempty

Proof let

stationary dist. x satisfies $x = Rx$

$$L = \{ \lambda > 0 \text{ s.t. } Ax \geq \lambda x \text{ for some } x \in \Delta^n \}$$

claims:

- L is nonempty
- L is bounded

if $\lambda \in L$, $Ax \geq \lambda x$ for $x \in \Delta^n$

then

$$\langle \mathbb{1}, A\mathbb{1} \rangle \geq \langle \mathbb{1}, Ax \rangle$$
$$\geq \langle \mathbb{1}, \lambda x \rangle = \lambda$$

"eigenvalue"

$$Ax = \lambda x$$

"eigenvector"

Lemma every linear map A has an eigenvalue $\lambda \in \mathbb{C}$ and eigenvector $x \in \mathbb{C}^n$

Lemma: A invertible $\Leftrightarrow$ A^T invertible

Lemma: A and A^T have same eigenvalues w/ same multiplicities

A is "**positive**" if $\langle x, Ay \rangle > 0$ $\forall$ nonzero $x, y \in \mathbb{R}^n_{\geq 0}$

Theorem: If A is positive, then A has eigenvalue λ_1, eigenvector x_1 s.t.

(a) $\lambda_1 > 0$ and $x_1 > 0$

(b) x_1 is unique eigenvector of λ_1

(c) x_1 is unique nonnegative vector s.t. $Ax_1 \geq \lambda_1 x_1$

(d) All other eigenvalues μ have $|\mu| < \lambda_1$

(e) No other eigenvectors are nonnegative

Proof let

stationary dist. x satisfies $x = Rx$

$$L = \{\lambda > 0 \text{ s.t. } Ax \geq \lambda x \text{ for some } x \in \Delta^n\}$$

"eigenvalue"

$$Ax = \lambda x$$

"eigenvector"

Lemma every linear map A has an eigenvalue $\lambda \in \mathbb{C}$ and eigenvector $x \in \mathbb{C}^n$

Lemma: A invertible $\iff A^T$ invertible

Lemma: A and A^T have same eigenvalues w/ same multiplicities

A is "positive" if $\langle x, Ay \rangle > 0 \ \forall$ nonzero $x, y \in \mathbb{R}^n_{\geq 0}$

claims:

- ✓ L is nonempty
- ✓ L is bounded
- L is closed

Theorem: If A is positive, then A has eigenvalue λ_1, eigenvector x_1 s.t.

(a) $\lambda_1 > 0$ and $x_1 > 0$

(b) x_1 is unique eigenvector of λ_1

(c) x_1 is unique nonnegative vector s.t. $Ax_1 \geq \lambda_1 x_1$

(d) All other eigenvalues μ have $|\mu| < \lambda_1$

(e) No other eigenvectors are nonnegative

Proof let

stationary dist. x satisfies $x = Rx$

$$L = \{\lambda > 0 \text{ s.t. } Ax \geq \lambda x \text{ for some } x \in \Delta^n\}$$

"eigenvalue"

$$Ax = \lambda x$$

"eigenvector"

claims: ✓ L is nonempty ✓ L is bounded

- L is closed

$\lambda_1, \lambda_2, \dots \in L \longrightarrow \bar{\lambda}$

need to show $\bar{\lambda} \in L$

let $x_i \in \Delta$ s.t. $Ax_i \geq \lambda_i x_i$

Δ^n compact $\Rightarrow$

subseq. $x_{i_1}, x_{i_2}, \dots \rightarrow \bar{x} \in \Delta$

Lemma every linear map A has an eigenvalue $\lambda \in \mathbb{C}$ and eigenvector $x \in \mathbb{C}^n$

Lemma:
A invertible $\Leftrightarrow$ A^T invertible

Lemma: A and A^T have same eigenvalues w/ same multiplicities

A is "positive" if $\langle x, Ay \rangle > 0$ $\forall$ nonzero $x, y \in \mathbb{R}^n_{\geq 0}$

Theorem: If A is positive, then A has eigenvalue λ_1, eigenvector x_1 s.t.

(a) $\lambda_1 > 0$ and $x_1 > 0$

(b) x_1 is unique eigenvector of λ_1

(c) x_1 is unique nonnegative vector s.t. $Ax_1 \geq \lambda_1 x_1$

(d) All other eigenvalues μ have $|\mu| < \lambda_1$

(e) No other eigenvectors are nonnegative

Proof let

$\boxed{\text{stationary dist. } x \text{ satisfies } x = Rx}$

$$L = \{\lambda > 0 \text{ s.t. } Ax \geq \lambda x \text{ for some } x \in \Delta^n\}$$

"eigenvalue"

$$Ax = \lambda x$$

"eigenvector"

- ✓ L is nonempty
- ✓ L is bounded
- ✓ L is closed

Lemma every linear map A has an eigenvalue $\lambda \in \mathbb{C}$ and eigenvector $x \in \mathbb{C}^n$

Lemma: A invertible $\Leftrightarrow$ A^T invertible

Lemma: A and A^T have same eigenvalues w/ same multiplicities

A is "positive" if $\langle x, Ay \rangle > 0$ $\forall$ nonzero $x, y \in \mathbb{R}^n_{\geq 0}$

Theorem: If A is positive, then A has eigenvalue λ_1, eigenvector x_1 s.t.

(a) $\lambda_1 > 0$ and $x_1 > 0$

(b) x_1 is unique eigenvector of λ_1

(c) x_1 is unique nonnegative vector s.t. $Ax_1 \geq \lambda_1 x_1$

(d) All other eigenvalues μ have $|\mu| < \lambda_1$

(e) No other eigenvectors are nonnegative

$\boxed{\text{stationary dist. } x \text{ satisfies } x = Rx}$

$$\underset{\text{"eigenvector"}}{A x} = \overset{\text{"eigenvalue"}}{\lambda} x$$

Lemma every linear map A has an eigenvalue $\lambda \in \mathbb{C}$ and eigenvector $x \in \mathbb{C}^n$

Lemma:
A invertible $\Leftrightarrow$ A^T invertible

Lemma: A and A^T have same eigenvalues w/ same multiplicities

A is "**positive**" if
$\langle x, Ay \rangle > 0$ $\forall$ nonzero $x, y \in \mathbb{R}^n_{\geq 0}$

Theorem: If A is positive, then A has eigenvalue λ_1, eigenvector x_1 s.t.

(a) $\lambda_1 > 0$ and $x_1 > 0$

(b) x_1 is unique eigenvector of λ_1

(c) x_1 is unique nonnegative vector s.t. $Ax_1 \geq \lambda_1 x_1$

(d) All other eigenvalues μ have $|\mu| < \lambda_1$

(e) No other eigenvectors are nonnegative

$\boxed{\text{stationary dist. } x \text{ satisfies } x = Rx}$

"eigenvalue"

$$Ax = \lambda x$$

"eigenvector"

Lemma every linear map A has an eigenvalue $\lambda \in \mathbb{C}$ and eigenvector $x \in \mathbb{C}^n$

Lemma:
A invertible <=> A^T invertible

Lemma: A and A^T have same eigenvalues w/ same multiplicities

A is "**positive**" if
$\langle x, Ay \rangle > 0 \ \forall$ nonzero $x, y \in \mathbb{R}^n_{\geq 0}$

Theorem: If A is positive, then A has eigenvalue λ_1, eigenvector x_1 s.t.

(a) $\lambda_1 > 0$ and $x_1 > 0$

(b) x_1 is unique eigenvector of λ_1

(c) x_1 is unique nonnegative vector s.t. $Ax_1 \geq \lambda_1 x_1$

(d) All other eigenvalues μ have $|\mu| < \lambda_1$

(e) No other eigenvectors are nonnegative

Perron-Frobenius for random walks

Let $R: \mathbb{R}^V \to \mathbb{R}^V$ be random walk map on strongly connected graph.

(a) $\exists$ eigenvector $x \in \Delta^n$ w/ eigenvalue 1

(b) x is unique eigenvector w/ eigenvalue 1

(c) x is only nonnegative eigenvector

(d) Other eigenvalues μ have $|\mu| < 1$

"eigenvalue"

$$Ax = \lambda x$$

"eigenvector"

Lemma every linear map A has an eigenvalue $\lambda \in \mathbb{C}$ and eigenvector $x \in \mathbb{C}^n$

Lemma: A invertible $\Leftrightarrow A^T$ invertible

Lemma: A and A^T have same eigenvalues w/ same multiplicities

A is "positive" if $\langle x, Ay \rangle > 0 \ \forall$ nonzero $x, y \in \mathbb{R}^n_{\geq 0}$

Theorem: If A is positive, then A has eigenvalue λ_1, eigenvector x_1 s.t.

(a) $\lambda_1 > 0$ and $x_1 > 0$

(b) x_1 is unique eigenvector of λ_1

(c) x_1 is unique nonnegative vector s.t. $Ax_1 \geq \lambda_1 x_1$

(d) All other eigenvalues μ have $|\mu| < \lambda_1$

(e) No other eigenvectors are nonnegative

Proof

$$Ax = \lambda x$$

"eigenvalue" "eigenvector"

Lemma every linear map A has an eigenvalue $\lambda \in \mathbb{C}$ and eigenvector $x \in \mathbb{C}^n$

Lemma:
A invertible $\iff A^T$ invertible

Lemma: A and A^T have same eigenvalues w/ same multiplicities

A is "**positive**" if
$\langle x, Ay \rangle > 0 \;\; \forall$ nonzero $x, y \in \mathbb{R}^n_{\geq 0}$

Theorem: If A is positive, then A has eigenvalue λ_1, eigenvector x_1 s.t.

(a) $\lambda_1 > 0$ and $x_1 > \mathbb{0}$

(b) x_1 is unique eigenvector of λ_1

(c) x_1 is unique nonnegative vector s.t. $Ax_1 \geq \lambda_1 x_1$

(d) All other eigenvalues μ have $|\mu| < \lambda_1$

(e) No other eigenvectors are nonnegative

Let $R: \mathbb{R}^V \to \mathbb{R}^V$ be random walk map on strongly connected graph.

(a) $\exists$ eigenvector $x \in \Delta^n$ w/ eigenvalue 1

(b) x is unique eigenvector w/ eigenvalue 1

(c) x is only nonnegative eigenvector

(d) Other eigenvalues μ have $|\mu| < 1$

Chapter 16

A Linear Map of the Web

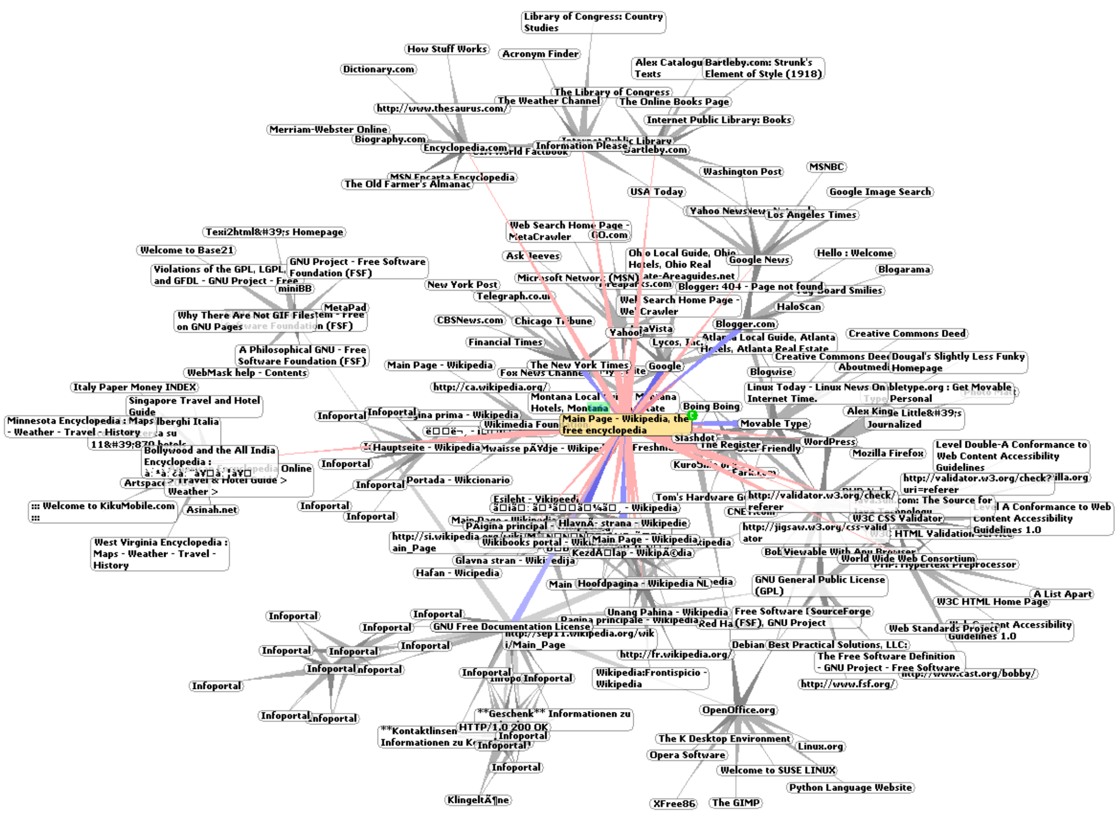


Figure 16.1: Map of a small part of the world wide web around wikipedia.org [CW].

16.1 Ranking the web

The world wide web is a large and messy place. As of March 2022, http://worldwidewebsize.com estimates that there are 2.97 *billion* indexed webpages. Even more amazing is that modern search engines, starting with Google, are able to process, organize and index this nearly unbounded corpus and make it useful. You can query for a topic of interest, and the search engine returns a long list of relevant websites, almost immediately. More often than not, you find what you are looking for within the first few listed results. This is utterly amazing, and we basically take it for granted.

It is one thing to identify all the web pages containing (or relevant to) a search query. This requires crawling the internet, and building a huge index that roughly identifies which keywords appear where. Some of the randomized data structures discussed earlier may be helpful for managing this task. But even within a search query, there seems to be an unlimited number of pages about a given topic, and a lot of it is junk. There is still another challenge to identify the best pages for the query. How do we separate the good websites from the bad? We should keep in mind the *scale* of the world wide web. It is pointless to try to evaluate the websites individually. This is a large scale *ranking* problem.

Modern search engines are based on the idea that the *link structure* of the world wide web reveals some sense of importance among the websites. When we write a paper, we cite the references that support or inform our argument. Likewise, web sites link to other websites and thereby implicitly bestow some approval. Another appeal of analyzing links is that we can model everything in basic graph theory, where we have good algorithms and sound analysis. The link structure gives a good starting point for our first idea for ranking webpages.

Idea 1. *Score each website equal to the number of other webpages linking to it.*

$$\mathrm{score}_1(v) = \sum_{(u,v)\in E} 1.$$

One feature of score_1 is that every link out of a vertex u is worth 1 point. But if u has many outgoing links, shouldn't that dilute the "approval" bestowed by u? As an analogy, suppose we have two lists of movies. One lists the top 10 movies of all time, and the other lists the top 100 movies of all time. Shouldn't it be worth more to be on the first list? Our next proposal for ranking scales down the value of a link (u, v) by the number of outgoing edges of u, so that the total sum of links leaving u is 1.

Let $d^+(u)$ denote the number of edges leaving a vertex u, a.k.a. the *out-degree*.

Idea 2. *Score each website v as the sum, over all other webpages u linking to v, of* 1 *divided by the number of outgoing links from u:*

$$\text{score}_2(v) = \sum_{u:(u,v)\in E} \frac{1}{d^+(u)}.$$

Unfortunately score_2 can be manipulated as follows. To promote a website v, one need only create many fake websites u with a link to v to drive up $\text{score}_2(v)$. This suggests that, when evaluating a link (u, v), we should consider whether u is much of an authority to begin with.

Idea 3. *Score each website v equal to the weighted sum, over all other webpages u linking to it, of the score of u divided by the number of outgoing links from u:*

$$\text{score}_3(v) = \sum_{u:(u,v)\in E} \frac{\text{score}_3(u)}{d^+(u)}.$$

$\text{score}_3(v)$ attain a sort of self-consistent nirvana. For example, the weight of a link (u, v), in attributing authority to v, is adjusted in proportion to the authority of u. The total authority distributed by a webpage u is exactly equal to u's own authority, $\text{score}_3(u)$.

While this recursive relationship is appealing, there is no reason, *a priori*, why such scores should exist.

Today's discussion is about how $\{\text{score}_3(v),\ v \in V\}$ *does* exist, and the structure and interpretations thereof. This happy miracle is entirely due to the fact that the values $\{\text{score}_3(v), v \in V\}$ satisfy a particularly well-structured *linear system of equations*. As such, our discussion will soon be translated into linear algebra. An important part of the structure comes from a probabilistic interpretation of the linear system.[1]

The goal of this discussion is to prove the following theorem (in more general terms). Recall that $\Delta^V = \left\{x \in \mathbb{R}^V_{\geq 0} : \sum_{v\in V} x_v = 1\right\}$ denotes the set of probability vectors over V.

Theorem 16.1. *There exists a vector $x \in \Delta^V$ such that*

$$x_v = \sum_{u:(u,v)\in E} \frac{x_u}{d^+(u)} \tag{16.1}$$

for all $v \in V$. If G is strongly connected, then this vector is unique and strictly positive.

[1] Naturally, the mathematics we discuss existed long before search engines and similar ideas had been applied elsewhere.

To try to understand where such an x comes from — and in particular, how we can assert that it describes a distribution over B — consider the following *random walk* on G. At each step, you are on some vertex u. You choose an outgoing edge $(u, v) \in E$ uniformly at random, and step to v. This may send you walking chaotically all over the graph. On the world wide web, this is like randomly surfing the web where you keep following randomly chosen links. If the links tend to point to useful websites, then over time your random walk should stumble upon good web sites more often then a uniformly random sample of the web. Now, depending on where you start, after some number of k steps, there are different probabilities of where you would end up. As k increases, we might hope this reaches some kind of equilibrium, where the distribution is the same or close to the same between to kth and $(k+1)$th step for sufficiently large k. This brings us to the notion of a stationary distribution.

Definition 16.2. *Fix a random walk on a set of vertices V. A set of probabilities $x \in \Delta^V$ is a* stationary distribution *for the random walk if taking a random step from the distribution x produces the same distribution x.*

Consider again theorem 16.1, and consider the recursive relations satisfied by $x \in \Delta^V$:

$$x_v = \sum_{u:(u,v)\in E} \frac{x_u}{d^+(u)}$$

for all $v \in V$. In our random walk, from a given vertex u, we choose an outgoing edge (u, v) with probability $1/d^+(u)$. In particular, if u is chosen from a random distribution $x \in \Delta^V$, then the probability of then stepping to a particular vertex v is

$$\sum_{u:(u,v)\in E} \frac{x_u}{d^+(u)}.$$

For the vector x asserted by theorem 16.1, this sum equals x_v. That is, *x is the stationary distribution of a random walk on G.* We can restate theorem 16.1 as follows.

Theorem 16.1, restated. *Every random walk on a directed graph G has a stationary distribution. If G is strongly connected, then this distribution is unique.*

16.1.1 PageRank

The actual PageRank algorithm proposed in [PBMW99] is slightly different. We augment the random walk thought experiment described above with a small probability α of restarting the walk from a page chosen uniformly at random. Otherwise (with the remaining probability $1 - \alpha$), we continue the random walk as described above. The same conclusion holds for this random walk.

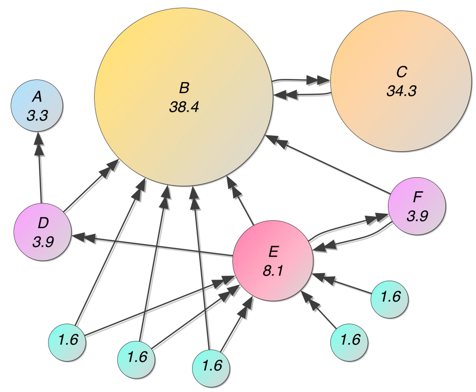


Figure 16.2: An example of PageRank on an 11 vertex graph [Wika].

Theorem 16.3. *Let $G = (V, E)$ be a directed graph, and let $\alpha \in (0, 1)$. There exists a unique vector $x \in \Delta^V$ such that for all $v \in V$,*

$$x_v = (1 - \alpha) \sum_{u:(u,v)\in E} \frac{x_u}{d^+(u)} + \frac{\alpha}{n}.$$

The PageRank formulation has some additional convenient properties compared to general random walks. First, the vector x is guaranteed to be unique, and has strictly positive coordinates. (Implicitly, it is the stationary distribution on the directed graph augmented by weighted edges between all pairs, which by theorem 16.1 is unique.) Second, it can be calculated more directly. We will come back to this point at the end of our discussion in section 16.6.

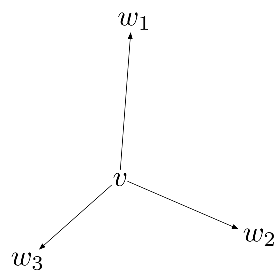
16.2 A linear map of the web

Recall that a function $f : \mathbb{R}^n \to \mathbb{R}^n$ is *linear* if it satisfies the following:

$$f(x + y) = f(x) + f(y)$$

Let $A : \mathbb{R}^V \to \mathbb{R}^V$ be the linear map encoding the directed edges as follows. For a vertex v, let $e_v \in \{0, 1\}^V$ be the vector with 0's everywhere except for 1 in the vth coordinate. We define A by setting $(Ae_v) \in \{0, 1\}^V$ to indicate the outgoing neighbors of v. More explicitly, A is defined by

$$\langle e_w, Ae_v\rangle = (Ae_v)_w = \begin{cases} 1 & \text{if } (v, w) \in E \\ 0 & \text{otherwise.} \end{cases}$$



$$\mapsto \quad (Ae_v) \in \{0,1\}^V \quad \begin{array}{c l} \left(\begin{array}{c} \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \end{array}\right) & \begin{array}{l} \\ \\ w_1 \\ \\ \\ \\ w_2 \\ \\ \\ \\ w_3 \\ \\ \\ \end{array} \end{array}$$

Figure 16.3: The adjacency vector (Ae_v) of a vertex v.

See also fig. 16.3. Note that the graph G is undirected iff A is symmetric.

More generally, we might have weights on the edges of the graph. Then we would define $A : \mathbb{R}^V \to \mathbb{R}^V$ by

$$\langle e_w, Ae_v \rangle = (Ae_v)_w = \begin{cases} \text{weight of the edge } (v,w) & \text{if } (v,w) \in E \\ 0 & \text{otherwise.} \end{cases}$$

Let us rehash our discussion on scoring websites in terms of the linear map. We have

$$\text{score}_1(v) = \langle e_v, A\mathbb{1} \rangle = \sum_w (Ae_v)_w.$$

We also have

$$d^+(u) = \langle \mathbb{1}, Ae_u \rangle.$$

Let $D = \text{diag}(d^+) \in \mathbb{R}^{V \times V}$ be the diagonal matrix induced by D. That is, for any vector x and vertex v, we have

$$(Dx)_v = d^+(v)x_v.$$

Then we can write $\text{score}_2(v)$ as

$$\text{score}_2(v) = \sum_{u:(u,v)\in E} \frac{1}{d^+(u)} = (AD^{-1}\mathbb{1})_v = \left\langle e_v, AD^{-1}\mathbb{1}\right\rangle.$$

As for the stationary distribution (score_3), we see that the stationary distribution x must satisfy the equation

$$x = Rx \text{ for the linear map } R = AD^{-1} : \mathbb{R}^V \to \mathbb{R}^V.$$

The map R has the property that $Rx \in \Delta^V$ for any $x \in \Delta^V$. Any linear map $A : \mathbb{R}^n \to \mathbb{R}^n$ mapping Δ^n to Δ^n is called a *stochastic* linear function.

16.3 Eigenvectors

Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map. An *eigenvector* of A is a nonzero vector $x \neq \mathbb{0}$ such that

$$Ax = \lambda x$$

for some scalar value $\lambda \in \mathbb{C}$. The scalar λ is called an *eigenvalue* of A, and here it is the eigenvalue corresponding to the eigenvector x.

Alternatively, a value $\lambda \in \mathbb{C}$ is an eigenvalue iff the linear map $(A - \lambda I) : \mathbb{R}^n \to \mathbb{R}^n$ is not invertible. Indeed, any eigenvector x corresponding to λ gives a second vector (besides $\mathbb{0}$) such that $(A - \lambda I)x = \mathbb{0}$.

Lemma 16.4. *The set of eigenvectors corresponding to an eigenvalue λ form a vector space.*

The dimension of the subspace corresponding to an eigenvalue λ is called the *multiplicity* of the eigenvalue. For any eigenvalue λ of $A : \mathbb{R}^n \to \mathbb{R}^n$, the multiplicity of λ is equal to $n - \text{rank}(A - \lambda I)$.

Existence of eigenvectors. *A priori*, it is not clear why every matrix should have an eigenvector.

Lemma 16.5. *Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map. Then A has an eigenvalue $\lambda \in \mathbb{C}$ and an eigenvector $x \in \mathbb{C}^n$.*

Proof. Fix any nonzero vector v. Recall that any set of $n+1$ vectors in $\mathbb{R}^n$ is linearly dependent. In particular, the set $v, Av, \dots, A^n v$ is linearly dependent. Put alternatively, there is a nonzero, degree $k \leq n$ polynomial $p(x) = \alpha_k x^k + \cdots + \alpha_0$ such that

$$p(A)v = 0.$$

By the fundamental theorem of algebra, the polynomial $p(x)$ can be expressed as

$$p(x) = (x - r_1)(x - r_2) \cdots (x - r_k)$$

where $r_1, \ldots, r_n \in \mathbb{C}$ are complex roots of $p(x)$. Then

$$p(A)v = (A - r_1 I)(A - r_2 I) \cdots (A - r_k I)v = 0$$

imples that some $A - r_i I$ maps a nonzero vector to $\mathbb{0}$. The corresponding root r_i is an eigenvalue. □

Tranposing and Eigenvalues The eigenvalues and eigenvectors of a matrix A and its transpose A^T are closely related. This is primarily because A and its transpose A^T have the same rank, and eigenvalues are ultimately concerned with values λ for which $A - \lambda I$ is not full rank.

Lemma 16.6. *Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map. Then A is invertible iff A^T is invertible.*

Proof. Suppose A is invertible. We claim that A^T is invertible with inverse $(A^{-1})^T$. Indeed, for any two points x, y, we have

$$\left\langle x, (A^{-1})^T A^T y \right\rangle = \left\langle A^{-1}x, Ay \right\rangle = \left\langle AA^{-1}x, y \right\rangle = \langle x, y \rangle.$$

Since this holds for all x and y, we have that $(A^{-1})^T A = I$. That is, $(A^{-1})^T$ is an inverse for A. This proves the "only if" whereas we claim "if and only if"; the "if" follows symmetrically as $(A^T)^T = A$. □

Lemma 16.7. *Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map.. Then A and A^T have the same eigenvalues with the same multiplicities.*

Proof. Recall that for any map $L : \mathbb{R}^n \to \mathbb{R}^n$, L is invertible iff L^T is invertible, and $(L^T)^{-1} = L^T$. Now suppose that λ is an eigenvalue of A. Then $A - \lambda I$ is not invertible, hence $(A - \lambda I)^T = A^T - \lambda I$ is not invertible, and so λ is an eigenvalue of A. The multiplicities are equal because the

$$\operatorname{rank}(A - \lambda I) = \operatorname{rank}\Big((A - \lambda I)^T\Big) = \operatorname{rank}(A - \lambda I).$$

□

Lemma 16.8. *Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map.. Let x be an eigenvector of A and let y be an eigenvector of A^T corresponding to distinct eigenvalues. The $\langle x, y \rangle = \mathbb{0}$.*

Proof. Let x have corresponding eigenvalue λ and let y have corresponding eigenvalue μ. We have

$$\lambda\langle x, y\rangle = \langle Ax, y\rangle = \left\langle x, A^T y\right\rangle = \mu\langle x, y\rangle.$$

If $\lambda \neq \mu$, then $\langle x, y \rangle = 0$. □

Let us now restate theorem 16.1 in our new language of eigenvectors and eigenvalues.

Theorem 16.1, in terms of eigenvectors and eigenvalues. *Let $G = (V, E)$ be a directed graph and let $R : \mathbb{R}^V \to \mathbb{R}^V$ be the map corresponding to the linear map corresponding to the random walk on G. Then R has an eigenvector $x \in \Delta^V$ with eigenvalue 1. If G is strongly connected, then x is the unique eigenvector (up to scaling) with eigenvalue 1.*

16.4 The Perron-Frobenius theorem

Definition 16.9. *Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map. A is* positive *if $\langle x, Ay \rangle > 0$ for all $x, y \in \mathbb{R}^n_{\geq 0}$ with $x, y \neq \mathbb{0}$.*

The following theorem is called the *Perron-Frobenius theorem.*

Theorem 16.10. *Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map such that $Ax > \mathbb{0}$ (coordinatewise) for all $x \neq \mathbb{R}^n_{\geq 0}$ with $x \neq \mathbb{0}$. Then A has an eigenvalue λ_1 with eigenvector x_1 with the following properties.*

1. *$\lambda_1 > 0$ and $x_1 > \mathbb{0}$.*
2. *x_1 is the unique (generalized) eigenvector of λ_1.*
3. *x_1 is also the unique nonnegative vector such that $Ax_1 \geq \lambda_1 x_1$ (up to scaling).*
4. *Any other eigenvalue μ has $|\mu| < x_1$.*[2]
5. *Any other eigenvector of A has at least one negative entry.*

[2] We will only prove that $|\mu| \leq x_1$.

Proof. Let

$$L = \{\lambda > 0 \text{ where } Ax \geq \lambda x \text{ for some } x \in \Delta^n\}.$$

We will argue that the the supremum of L is the desired value λ_1. To this end, we first make the following claims about L.

1. *L is nonempty, and contains a positive number.*
2. *L is bounded.*
3. *L is closed.*

For claim 1, let $x \in \Delta^n$ with $x > \mathbb{0}$. The Ax will have strictly positive coordinates because the graph is strongly connected, and so there is some $\lambda > 0$ such that $Ax > \lambda x$.

For claim 2, we observe that for $\lambda \in L$, with (say) $Ax \geq \lambda x$ with $x \in \Delta_n$, we have

$$\langle \mathbb{1}, A\mathbb{1} \rangle \overset{\text{(a)}}{\geq} \langle \mathbb{1}, Ax \rangle \overset{\text{(b)}}{\geq} \langle \mathbb{1}, \lambda x \rangle \overset{\text{(c)}}{=} \lambda.$$

(a) is because A is monotonic and $\mathbb{1} \geq x$. (b) is by choice of x and λ. (c) is because $x \in \Delta_n$.

For claim 3, let $\lambda_1, \lambda_2, \ldots$ be any sequence of points in L that converges to some $\bar{\lambda}$. We want to show that $\bar{\lambda} \in L$. For each i, let $x_i \in \Delta^n$ with $Ax_i \geq \lambda_i x_i$. Since Δ^n converges, the x_i's converge to some $\bar{x} \in \Delta^n$. We have

$$A\bar{x} = A\Big(\lim_{n\to\infty} x_i\Big) \overset{\text{(d)}}{=} \lim_{n\to\infty} Ax_i \geq \lim_{n\to\infty} \lambda_i x_i = \lambda\bar{x}.$$

(d) invokes continuity of A to pass through the limit.

We have now show that L is a closed and bounded set with at least one positive number. Any nonempty closed and bounded set L *has a finite supremum, λ_1, which is contained in L.* By definition of L, there is also a vector $x_1 \in \Delta^n$ such that $Ax_1 \geq \lambda_1 x_1$. *We claim that $Ax_1 = \lambda_1 x_1$.*

Let $y \geq \mathbb{0}$ be such that $Ax_1 = \lambda_1(x_1 + y)$. Then $Ax_1 = \lambda_1 x_1 \iff y = \mathbb{0}$. Suppose by contradiction that $y \neq \mathbb{0}$. Choose $\epsilon \in (0, 1)$ sufficiently small that

$$\frac{\epsilon}{1 - \epsilon} \leq \langle \mathbb{1}, y \rangle.$$

Observe that

$$(1 - \epsilon)x_1 + \frac{\epsilon}{\langle \mathbb{1}, y \rangle} \in \Delta^n.$$

Moreover, we have

$$A\left((1-\epsilon)x_1 + \frac{\epsilon}{\langle \mathbb{1}, y\rangle} y\right) \overset{\text{(e)}}{>} (1-\epsilon)Ax_1 \overset{\text{(f)}}{=} (1-\epsilon)\lambda(x_1+y) \overset{\text{(g)}}{\geq} \lambda\left((1-\epsilon)x_1 + \frac{\epsilon}{\langle \mathbb{1}, y\rangle} y\right).$$

Here (e) is because $Ax > \mathbb{0}$ for all $x \geq \mathbb{0}$ with $x \neq \mathbb{0}$. (f) is by choice of y. (g) is by choice of ϵ. The strict inequality obtained above implies that there is a larger value than λ in L, a contradiction. Thus we must have $Ax_1 = A\lambda_1$ after all.

Note also that x_1 is strictly positive, as $x_1 = \lambda^{-1}Ax_1 > \mathbb{0}$. More generally, any eigenvector of A associated with a positive eigenvalue is strictly positive.

Next we claim that x_1 is the unique (simple) eigenvector for λ_1 (up to scaling). Indeed, suppose $Ay = \lambda_1 y$ for vector y. As mentioned above, we must have $y > \mathbb{0}$. If y is not proportional to x, then let $z = x - \alpha y$ where $\alpha > 0$ is such that $z \geq 0$, $z \neq \mathbb{0}$, and $z_i = 0$ for some coordinate i. Then z would be an eigenvector of A that is not strictly positive, a contradiction.

Now we claim that x_1 is the unique generalized eigenvector for λ_1, as well. If not, then there would a vector y not spanned by x_1 such that

$$(A - \lambda_1 I)^2 y = 0.$$

Then $(A - \lambda_1 I)y$ is a simple eigenvector of A with eigenvalue λ_1, so

$$Ay = \lambda_1 y + cx_1 \tag{16.2}$$

for some $c \neq 0$. By flipping the sign of y if necessary, we may assume that $c > 0$. Increase y by a multiple of x_1 if necessary, we may assume that $y > 0$. If $y > 0$ and $c > 0$, then (2) that there is a value larger than λ_1 in L, a contradiction. Thus x_1 is the unique generalized eigenvector for λ_1.

Take any other eigenvalue μ of A, with eigenvector y. Scale y such that Then

$$|\mu||y| = |\mu y| = |Ay| \leq A|y|,$$

where $|y|$ denotes the coordinate-wise absolute value of y. Thus $|\mu| \in L$, and therefore $|\mu| \leq x_1$.

If $|\mu| = \lambda$, then we would have $|y| = x_1$ (after scaling) by uniqueness of x_1. This, combined with $|Ay| = A|y|$, implies (by known facts about complex numbers, which we omit) that $y = e^{i\theta}x_1$ for some fixed θ. Thus $y \in \mathrm{span}(x_1)$ and $\mu = x_1$.

For the last claim, observe that A^T also satisfies the hypothesis. Indeed, if $x, y \geq \mathbb{0}$ and neither x nor y equals zero, then

$$\left\langle A^T x, y\right\rangle = \langle x, Ay\rangle > 0.$$

Thus A^T has (the same) dominant eigenvalue λ_1, with a positive eigenvector $y_1 \in \mathbb{R}^n_{>0}$. Now, any eigenvector x of A corresponding to an eigenvalue other than x_1 must be orthogonal to y_1. If y_1 is has strictly positive coordinates and $\langle x, y_1 \rangle = 0$, then x must have at least one negative coordinate. □

The above proof is essentially due to Bohnenblust; see [Bel97; Lax07].

16.5 Perron-Frobenius for strongly connected random walks

We now extend theorem 16.10 to random walks. In particular, we will show that there is always a stationary distribution, and that this stationary distribution is unique if the underlying graph is strongly connected.

Theorem 16.11. *Let $A : \mathbb{R}^V \to \mathbb{R}^V$ be the linear map of a random walk on a strongly connected graph.*

1. *There is an eigenvector $x \in \Delta^n$ with eigenvalue 1 and $x > \mathbb{0}$.*
2. *x is the unique eigenvector with eigenvalue 1.*
3. *x is the only eigenvector of A with no negative entries.*
4. *Any other eigenvalue μ has $|\mu| \leq 1$.*

Proof. We first note that A has eigenvalue 1. Indeed, because A models a random walk, we have

$$A^T \mathbb{1} = \mathbb{1},$$

as can be verified directly. Thus 1 is an eigenvalue of A^T and thereby an eigenvalue of A as well. Now, let x be any eigenvector of A, rescaled so that $\langle \mathbb{1}, x \rangle = 1$. We claim that $x \in \Delta^n$.

Consider the matrix $B = \frac{1}{n} \sum_{i=0}^{n} A^i$. We can interpret B as the random walk on V induced by the following two steps:

1. Choose $i \in \{0, \ldots, n-1\}$ uniformly at random.
2. Take i steps of the random walk A.

Observe that any eigenvalue μ of A corresponds to an eigenvalue of B with value

$$\frac{1}{n} \sum_{i=0}^{n-1} \mu^i,$$

with the same eigenvectors. In particular, 1 is an eigenvector of x with eigenvalue 1.

Because G is strongly connected, there is a path of at most $n-1$ edges from any point $a \in V$ to any point $b \in V$. Thus B models a random walk with strictly positive transition probabilities for all pairs of vertices. In particular, B satisfies the positive assumptions of theorem 16.10.

Let $\lambda_1 > 0$ and $x_1 \in \Delta^n$ be the "dominant" eigenvector and eigenvalue. We claim that $\lambda_1 = 1$ and $x_1 = x$. To this end, consider B^T. B^T has the same dominant eigenvalue λ_1, with a corresponding eigenvector that is also the only eigenvector with no negative corodinates. We also know that 1 is an eigenvalue of B^T with eigenvector $\mathbb{1}$. So $\lambda_1 = 1$. Theorem 16.10 then implies that x is the unique eigenvector for eigenvalue 1, and all the coordinates of x are strictly positive.

Let μ be any other eigenvalue of A. We claim that $|\mu| \leq 1$. For $\epsilon \geq 0$, let $A_\epsilon = (1-\epsilon)A + \epsilon B$. Let

$$\mu_\epsilon = \mu + \frac{\epsilon}{n}\sum_{i=1}^{n-1} \mu^i.$$

μ_ϵ is an eigenvalue of A_ϵ for all $\epsilon \geq 0$. Note that for all $\epsilon > 0$, theorem 16.10 applies to A_ϵ and in particular A_ϵ has dominant eigenvalue 1, and $|\mu_\epsilon| \leq 1$. Moreover, $\lim_{\epsilon \to 0} \mu_\epsilon = \mu$. Thus $|\mu| \leq 1$.

Any eigenvector other than x is an non-dominant eigenvector of B, and thus has negative coordinates. □

16.6 Computing PageRank

The PageRank vector x satisfies the equation

$$x = (1-\epsilon)Rx + \frac{\epsilon}{n}\mathbb{1},$$

where $R = AD^{-1}$ models the random walk on the directed graph, and where $\mathbb{1} \in \mathbb{R}^V$ is the all-1's vector. We can rewrite this as

$$\left(\left(\frac{1}{1-\epsilon}\right)I - R\right)x = \frac{\epsilon}{(1-\epsilon)n}\mathbb{1}.$$

Recall that the maximum eigenvalue of R is 1; in paricular, $1/(1-\epsilon)$ is not an eigenvalue. Thus $(1/(1-\epsilon)I - R)$ is invertible. We have

$$x = \left(\frac{\epsilon}{(1-\epsilon)n}\right)\left(\frac{1}{1-\epsilon}I - R\right)^{-1}\mathbb{1} = \frac{\epsilon}{n}(I - (1-\epsilon)R)^{-1}\mathbb{1}.$$

We can write $(I - (1-\epsilon)R)^{-1}$ as the infinite series

$$(I - (1-\epsilon)R)^{-1} = \lim_{k\to\infty} \sum_{i=0}^{k} ((1-\epsilon)R)^i.$$

Thus

$$x = \frac{\epsilon}{(1-\epsilon)n} \lim_{k\to\infty} \sum_{i=0}^{k} ((1-\epsilon)R)^i \mathbb{1}.$$

The series on the RHS converges quickly for moderate ϵ, so in practice one only has to compute a few terms in the sum.

16.7 Exercises

Exercise 16.1. Let $G = (V, E)$ be a directed graph, not necessarily strongly connected. Recall that the strongly connected componentsm for a directed acyclic graph. A *sink component* is a strongly connected component with no out going edges; i.e., a sink in the DAG of strongly connected components. Suppose G has a unique sink component $S \subset V$. Show that the random walk on G has a unique stationary distribution $x \in \Delta^V$ and that $x_v > 0$ iff $v \in S$.

Exercise 16.2. Let $G = (V, E)$ be an simple[3], unweighted, directed, and strongly connected graph. We proved in theorem 16.1 that the random walk G has a unique and strictly positive stationary distribution $x \in \Delta^V$. Prove that for all $v \in V$, $x_v \geq n^{-(n+1)}$.

[3] *Simple* means that there are no parallel edges.