# Random Walks, PageRank, and Computer Science

Siu On Chan

Department of Computer Science and Engineering
The Chinese University of Hong Kong

July 30, 2015

# Yahoo in 1996: Human-edited directory

YAHOO!

NEW — COOL — RANDOM — HEAD LINES — YAHOO INFO — ADD URL

Yahoo! Deutschland | CLICK HERE TO VISIT THE STARS YAHOO! LOS ANGELES | Weekly Picks

[ ] Search Options

Yellow Pages - People Search - City Maps -- News Headlines - Stock Quotes - Sports Scores

- **Arts** - - *Humanities, Photography, Architecture, ...*
- **Business and Economy [Xtra!]** - - *Directory, Investments, Classifieds, ...*
- **Computers and Internet [Xtra!]** - - *Internet, WWW, Software, Multimedia, ...*
- **Education** - - *Universities, K-12, Courses, ...*
- **Entertainment [Xtra!]** - - *TV, Movies, Music, Magazines, ...*
- **Government** - - *Politics **[Xtra!]**, Agencies, Law, Military, ...*
- **Health [Xtra!]** - - *Medicine, Drugs, Diseases, Fitness, ...*
- **News [Xtra!]** - - *World **[Xtra!]**, Daily, Current Events, ...*

# Only 5 directory entries for “Linear Algebra”

**Top:Science:Mathematics:Linear Algebra**

Search Options

Search all of Yahoo Search only in **Linear Algebra**

- I Hate Linear Algebra
- Java Script Linear Algebra - calculate simple matrix operations.
- Linear Algebra WebNotes - undergraduate course at the University of Nebraska-Lincoln. Contains lectures, homework assignments, Maple examples, and a discussion page.
- Slitex Foilset CPS615 Gauss Seidel Finite Element Methods and Conjugate Gradient - discusses sequential and parallel Gauss Seidel and Jacobi iteration, finite element methods applied to Laplace's equation in 2D, and the Conjugate Gradient method.
- Solving of linear equationsystems

# Google today

- ~10,000,000 Linear Algebra pages, ranked mostly by importance

Linear algebra | Khan Academy
https://www.khanacademy.org/math/**linear-algebra** ▾ Khan Academy ▾
**Linear algebra** describes things in two dimensions, but many of the concepts can be extended into three, four or more. **Linear algebra** implies two dimensional ...
Vectors and spaces - Matrix transformations - Alternate coordinate systems - Vectors

Linear algebra - Wikipedia, the free encyclopedia
https://en.wikipedia.org/wiki/**Linear_algebra** ▾ Wikipedia ▾
**Linear algebra** is the branch of mathematics concerning vector spaces and **linear** mappings between such spaces. It includes the study of lines, planes, and subspaces, but is also concerned with properties common to all vector spaces.
List of linear algebra topics - Basis - Rank - Category:Linear algebra

Free Linear Algebra textbook
joshua.smcvt.edu/**linearalgebra**/ ▾ Saint Michael's College ▾
**Linear Algebra** is a text for a first US undergraduate **Linear Algebra** course. It is Free. You can use it as a main text, as a supplement, or for independent study.

Linear Algebra | Mathematics | MIT OpenCourseWare
ocw.mit.edu › Courses › Mathematics ▾ MIT OpenCourseWare ▾
This is a basic subject on matrix theory and **linear algebra**. Emphasis is given to topics that will be useful in other disciplines, including systems of equations, ...

Part I:

PageRank and Random walk

# PageRank Algorithm

- Named after Larry Page, Google cofounder and current CEO
- Determines webpages' importance only by link structure of the directed graph of webpages

## PageRank Algorithm

- Named after Larry Page, Google cofounder and current CEO
- Determines webpages' importance only by link structure of the directed graph of webpages
- Based on stationary distribution of a random walk

# PageRank Algorithm

- Named after Larry Page, Google cofounder and current CEO
- Determines webpages' importance only by link structure of the directed graph of webpages
- Based on stationary distribution of a random walk

- Spectral graph theory!
- For simplicity, focus on undirected, regular graphs in this talk

# Random Walk

- Stochastic progress on a graph (undirected for now)

# Random Walk

- Stochastic progress on a graph (undirected for now)
- Starts from a vertex, at each time step $t$ moves to a uniformly random neighbour of the current vertex

# Distribution $\mathbf{p}_t$

Probability transition matrix $\mathbf{K}$

Row vector $\mathbf{p}_t$: distribution at time $t$

- $\mathbf{p}_{t+1} = \mathbf{p}_t\mathbf{K}$
- $\mathbf{p}_t = \mathbf{p}_0\mathbf{K}^t$
- Does $\mathbf{p}_t$ converge as $t$ increases?

# Stationary distribution

If $\mathbf{p}_t$ converges, the limiting distribution $\mathbf{p}_\infty$ must be stationary

$$\mathbf{p}_\infty \mathbf{K} = \mathbf{p}_\infty$$

- uniform distribution always a stationary distribution (for an undirected, regular graph)

# Limiting distribution: unique?

# Limiting distribution: unique?

- Not unique on disconnected graphs
  (Some vertex not reachable from some other vertex via intermediate vertices)

# Limiting distribution: unique?

- Not unique on disconnected graphs
  (Some vertex not reachable from some other vertex via intermediate vertices)
- Not unique on bipartite graphs
  (Can partition all vertices into two subsets $V_1, V_2$ so that all edges only go between $V_1$ and $V_2$)

# Unique limiting distribution

### Theorem (Uniqueness)

*An undirected, regular, connected, non-bipartite graph has a unique stationary distribution* $\mathbf{p}_*$*.*
*Further, given any initial distribution* $\mathbf{p}_0$*,*

$$\lim_{t \to \infty} \mathbf{p}_t = \mathbf{p}_*.$$

# Eigenvalues and eigenvectors

$\lambda \in \mathbb{R}$ is an eigenvalue and $\mathbf{q} \in \mathbb{R}^n$ is an eigenvector of $\mathbf{K}$ if

$$\mathbf{q}\mathbf{K} = \lambda \mathbf{q}$$

Fact: Transition matrix $\mathbf{K}$ of an undirected graph is real symmetric

## Lemma (Spectral Theorem)

*Any $n \times n$ real symmetric matrix $\mathbf{K}$ has $n$ eigenvalue-eigenvector pairs*

$$\mathbf{q}^{(1)}\mathbf{K} = \lambda_1 \mathbf{q}^{(1)}$$

$$\vdots$$

$$\mathbf{q}^{(n)}\mathbf{K} = \lambda_n \mathbf{q}^{(n)}$$

*such that $\{\mathbf{q}^{(1)}, \ldots, \mathbf{q}^{(n)}\}$ is an orthogonal basis*

# Proof

## Theorem (Uniqueness)

*An undirected, regular, connected, non-bipartite graph has a unique stationary distribution* $\mathbf{p}_*$.
*Further, given any initial distribution* $\mathbf{p}_0$,

$$\lim_{t\to\infty} \mathbf{p}_t = \mathbf{p}_*.$$

## Proof.

By spectral theorem, $\{\mathbf{q^{(1)}}, \ldots, \mathbf{q^{(n)}}\}$ forms a basis. Expand

$$\mathbf{p}_0 = \sum_{i=1}^{n} \alpha_i \mathbf{q^{(i)}}.$$

$$\mathbf{p}_t = \mathbf{p}_0 \mathbf{K}^t = \sum_{i=1}^{n} \alpha_i \mathbf{q^{(i)}} \mathbf{K}^t.$$

# Proof (continued)

$$\mathbf{p}_t = \mathbf{p}_0\mathbf{K}^t = \sum_{i=1}^{n} \alpha_i \mathbf{q}^{(\mathbf{i})}\mathbf{K}^t$$

Since

$$\mathbf{q}^{(\mathbf{i})}\mathbf{K}^t = \lambda_i \mathbf{q}^{(\mathbf{i})}\mathbf{K}^{t-1} = {\lambda_i}^2 \mathbf{q}^{(\mathbf{i})}\mathbf{K}^{t-2} = \cdots = {\lambda_i}^t \mathbf{q}^{(\mathbf{i})},$$

the top equation becomes

$$\mathbf{p}_t = \sum_{i=1}^{n} \alpha_i {\lambda_i}^t \mathbf{q}^{(\mathbf{i})}.$$

# Graph spectrum

Assume eigenvalues are sorted

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$$

# Graph spectrum

Assume eigenvalues are sorted

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$$

One can show

$$1 \geqslant \lambda_1 \qquad \text{and} \qquad \lambda_n \geqslant -1$$

Recall: $\lambda_1 = 1$, uniform distribution as an eigenvector

# Graph spectrum

Assume eigenvalues are sorted

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$$

One can show

$$1 \geqslant \lambda_1 \qquad \text{and} \qquad \lambda_n \geqslant -1$$

Recall: $\lambda_1 = 1$, uniform distribution as an eigenvector

## Proposition

$\lambda_2 = 1$ *if and only if disconnected graph*

# Graph spectrum

Assume eigenvalues are sorted

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$$

One can show

$$1 \geqslant \lambda_1 \qquad \text{and} \qquad \lambda_n \geqslant -1$$

Recall: $\lambda_1 = 1$, uniform distribution as an eigenvector

### Proposition

$\lambda_2 = 1$ *if and only if disconnected graph*

### Proposition

$\lambda_n = -1$ *if and only if bipartite graph*

# Proof (final bits)

$$\mathbf{p}_t = \sum_{i=1}^{n} \alpha_i {\lambda_i}^t \mathbf{q}^{(\mathbf{i})}.$$

For regular, connected, bipartite graph, $|\lambda_2| < 1, \ldots, |\lambda_n| < 1$. Hence $\mathbf{p}_t \to \alpha_1 \mathbf{q}^{(\mathbf{1})}$.

We are done if $\alpha_1 = 1$. To determine $\alpha_1$, recall $\mathbf{p}_0 = \sum_{i=1}^{n} \alpha_i \mathbf{q}^{(\mathbf{i})}$. Taking inner product with $\mathbf{q}^{(\mathbf{1})}$:

$$\mathbf{p}_0 \cdot \mathbf{q}^{(\mathbf{1})} = \sum_{i=1}^{n} \alpha_i \mathbf{q}^{(\mathbf{i})} \cdot \mathbf{q}^{(\mathbf{1})} = \alpha_1 \mathbf{q}^{(\mathbf{1})} \cdot \mathbf{q}^{(\mathbf{1})}.$$

Since $\mathbf{q}^{(\mathbf{1})}$ must be (a scalar multiple of) the uniform distribution,

$$\mathbf{p}_0 \cdot \mathbf{q}^{(\mathbf{1})} = 1/n \qquad \text{and} \qquad \mathbf{q}^{(\mathbf{1})} \cdot \mathbf{q}^{(\mathbf{1})} = 1/n,$$

hence $\alpha_1$ must be $1$. $\square$

# Directed graphs and PageRank

### Theorem (Uniqueness)

*An undirected, regular, connected, non-bipartite graph has a unique stationary distribution* $\mathbf{p}_*$.
*Further, given any initial distribution* $\mathbf{p}_0$,

$$\lim_{t\to\infty} \mathbf{p}_t = \mathbf{p}_*.$$

- A similar theorem (suitably modified) holds for directed, non-regular graphs: Perron–Frobenius theorem
- Limiting distribution $\mathbf{p}_*$ not necessarily uniform
- PageRank iteratively computes the distribution $\mathbf{p}_t = \mathbf{p}_0\mathbf{K}^t$ from an arbitrary initial distribution $\mathbf{p}_0$

Part II:

Connections to Theoretical Computer Science

# Spectral graph theory and expanders

- Spectral graph theory: study of graph eigenvalues $\lambda_1, \ldots, \lambda_n$ and graph properties

# Spectral graph theory and expanders

- Spectral graph theory: study of graph eigenvalues $\lambda_1, \ldots, \lambda_n$ and graph properties
- Graphs with $\lambda_2$ much smaller than $\lambda_1 = 1$ are called expanders
  Valuable to computer science

# Spectral graph theory and expanders

- Spectral graph theory: study of graph eigenvalues $\lambda_1, \ldots, \lambda_n$ and graph properties
- Graphs with $\lambda_2$ much smaller than $\lambda_1 = 1$ are called expanders
  Valuable to computer science
- For $d$-regular graphs, how small can $\lambda_2$ be?
  Recent breakthrough: Yale theoretical computer scientists (Marcus, Spielman, and Srivastava) constructed bipartite graphs for any degree $d$ with
  $$\max\{|\lambda_2|, |\lambda_{n-1}|\} \leqslant 2\sqrt{d-1}/d.$$
  Smallest possible (Alon–Boppana)

# Spectral graph theory and expanders

- Spectral graph theory: study of graph eigenvalues $\lambda_1, \ldots, \lambda_n$ and graph properties
- Graphs with $\lambda_2$ much smaller than $\lambda_1 = 1$ are called expanders
  Valuable to computer science
- For $d$-regular graphs, how small can $\lambda_2$ be?
  Recent breakthrough: Yale theoretical computer scientists (Marcus, Spielman, and Srivastava) constructed bipartite graphs for any degree $d$ with
  $$\max\{|\lambda_2|, |\lambda_{n-1}|\} \leqslant 2\sqrt{d-1}/d.$$
  Smallest possible (Alon–Boppana)
  Their novel techniques also resolve 54-year-old Kadison–Singer problem in Mathematics and engineering

# Matrix multiplication and computational complexity

- Given two matrices $\mathbf{A}$ and $\mathbf{B}$ of size $n$, compute $\mathbf{AB}$
- Recall $(\mathbf{AB})_{ij} = \sum_k \mathbf{A}_{ik}\mathbf{B}_{kj}$
  Straightforward algorithm requires roughly $n^3$ elementary operations

# Matrix multiplication and computational complexity

- Given two matrices $\mathbf{A}$ and $\mathbf{B}$ of size $n$, compute $\mathbf{AB}$
- Recall $(\mathbf{AB})_{ij} = \sum_k \mathbf{A}_{ik}\mathbf{B}_{kj}$
  Straightforward algorithm requires roughly $n^3$ elementary operations
- Strassen algorithm: roughly $n^{\log_2 7} \approx n^{2.807}$ elementary operations
- Le Gall algorithm (current best): roughly $n^{2.373}$ elementary operations

# Matrix multiplication and computational complexity

- Given two matrices $\mathbf{A}$ and $\mathbf{B}$ of size $n$, compute $\mathbf{AB}$
- Recall $(\mathbf{AB})_{ij} = \sum_k \mathbf{A}_{ik}\mathbf{B}_{kj}$
  Straightforward algorithm requires roughly $n^3$ elementary operations
- Strassen algorithm: roughly $n^{\log_2 7} \approx n^{2.807}$ elementary operations
- Le Gall algorithm (current best): roughly $n^{2.373}$ elementary operations
- Is $n^2$ possible?
  If so, potentially very useful
  If not, why not?