• Instructor: He Wang Email: he.wang@northeastern.edu

Google's PageRank Algorithm

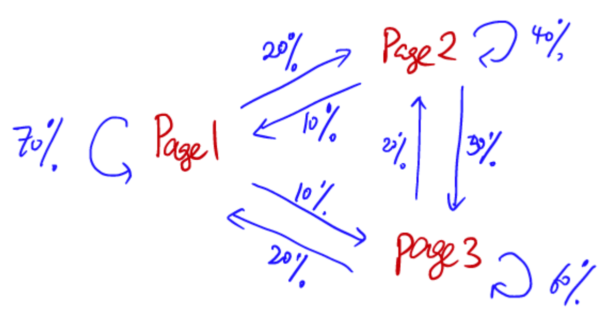
Google claims to index 25 billion pages. When you search a key word, how the search engine return pages in a good order.

The fundamental idea is created Google's founders, Sergey Brin and Lawrence Page: the importance of a page is judged by the number of pages linking to it as well as their importance.

Consider a mini-web with only three pages: Page1, Page2, Page3. Initially, there is an equal number of surfers on each page. The initial probability distribution vector is

$$\vec{x}_0 = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

After 1 minute, some people will move onto different pages with a probability distribution vector $\vec{x}_1$, as in the following diagram



They way can be described by a transformation matrix

$$A = \begin{bmatrix} 0.7 & 0.1 & 0.2 \\ 0.2 & 0.4 & 0.2 \\ 0.1 & 0.5 & 0.6 \end{bmatrix}$$

And we have

$$\vec{x}_1 = A\vec{x}_0$$

After another 1 minute, some people will move onto different pages with a probability distribution vector $\vec{x}_2$, such that

$$\vec{x}_2 = A\vec{x}_1 = A^2\vec{x}_0$$

After t minutes, probability distribution vector is

$$\vec{x}_t = A^t\vec{x}_0$$

2. Dynamical Systems and Eigenvectors.

Consider a sequence of linear transformations

$$\vec{x}(t+1) = A\vec{x}(t)$$

Each vector $\vec{x}(t)$ is called a **state vector**. Suppose we know the initial vector $\vec{x}(0) = \vec{x}_0$. We wish to find each state $\vec{x}(t)$:

$$\vec{x}(0) \xrightarrow{A} \vec{x}(1) \xrightarrow{A} \vec{x}(2) \xrightarrow{A} \cdots \xrightarrow{A} \vec{x}(t) \xrightarrow{A} \vec{x}(t+1) \xrightarrow{A} \cdots$$

That is

$$\vec{x}(t) = A^t\vec{x}(0) = A^t\vec{x}_0$$

Theorem.

Consider a **dynamical system**

$$\vec{x}(t+1) = A\vec{x}(t) \text{ with } \vec{x}(0) = \vec{x}_0.$$

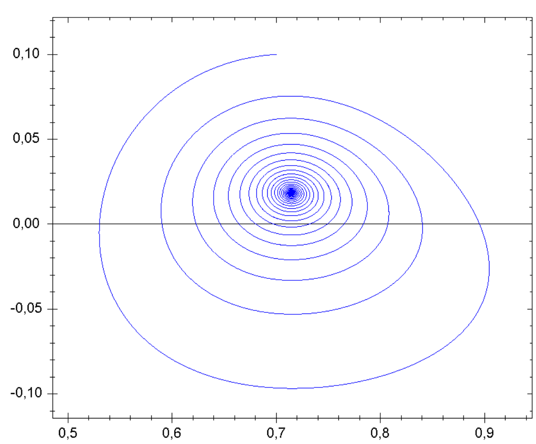
Then

$$\vec{x}(t) = A^t\vec{x}_0$$

Suppose A has a eigenbasis $\vec{b}_1, \ldots, \vec{b}_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$. Suppose $\vec{x}_0 = c_1\vec{b}_1 + c_2\vec{b}_2 + \cdots + c_n\vec{b}_n$. Then,

$$\begin{aligned}\vec{x}(t) &= A^t\vec{x}_0 \\ &= c_1A^t\vec{b}_1 + \cdots c_nA^t\vec{b}_n \\ &= c_1\lambda_1^t\vec{b}_1 + \cdots c_n\lambda_n^t\vec{b}_n\end{aligned}$$

Remark: Let A be a 2×2 matrix The endpoints of state vectors $\vec{x}(0)$, $\vec{x}(1)$, $\cdots$, $\vec{x}(t)$, $\ldots$, form the discrete **trajectory** of the system. A **phase portrait** of the dynamical system shows trajectories for various initial states.



Example 1.

$$A = \begin{bmatrix} 0.7 & 0.1 & 0.2 \\ 0.2 & 0.4 & 0.2 \\ 0.1 & 0.5 & 0.6 \end{bmatrix} \quad \text{and} \quad \vec{x}_0 = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

Find explicit formulas for A^t.

1. Find all eigenvalues by $\det(A-\lambda I)=0$

$\lambda_1 = 1, \quad \lambda_2 = 0.5, \quad \lambda_3 = 0.2$

2. Find an eigenbasis for A.

A basis for eigenspace E_{λ_1} is $\left\{ \begin{bmatrix} 7 \\ 5 \\ 8 \end{bmatrix} \right\}$

A basis for eigenspace E_{λ_2} is $\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}$

A basis for eigenspace $E_{\lambda_3} = \left\{ \begin{bmatrix} -1 \\ -3 \\ 4 \end{bmatrix} \right\}$

3. $A = PDP^{-1}$ where $P = \begin{bmatrix} 7 & 1 & -1 \\ 5 & 0 & -3 \\ 8 & -1 & 4 \end{bmatrix}$ $D = \begin{bmatrix} 1 & & \\ & 0.5 & \\ & & 0.2 \end{bmatrix}$

4. $A^t = PD^tP^{-1} = \begin{bmatrix} 7 & 1 & -1 \\ 5 & 0 & -3 \\ 8 & -1 & 4 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 0.5^t & \\ & & 0.2^t \end{bmatrix} \frac{1}{60} \begin{bmatrix} 3 & 3 & 3 \\ 44 & -16 & -16 \\ 5 & -15 & 5 \end{bmatrix}$

$= \begin{bmatrix} \vec{b}_1 & 0.5^t \vec{b}_2 & 0.2^t \vec{b}_3 \end{bmatrix} P^{-1}$

Example 2. Find explicit formulas for $A^t\vec{x}_0$

$[\vec{x}_0]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1/20 \\ -2/45 \\ -1/36 \end{bmatrix}$ $\quad \vec{x}_0 = c_1\vec{b}_1 + c_2\vec{b}_2 + c_3\vec{b}_3$ $\quad \vec{x}_0 = P\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$

$A^t\vec{x}_0 = c_1\lambda_1^t\vec{b}_1 + c_2\lambda_2^t\vec{b}_2 + c_3\lambda_3^t\vec{b}_3$

$= \frac{1}{20}\cdot\vec{b}_1 - \frac{2}{45}(0.5^t)\vec{b}_2 - \frac{1}{36}(0.2^t)\vec{b}_3$

Example 3. Find $\lim_{t\to\infty} A^t$

$$\lim_{t\to\infty} A^t = [\vec{b}_1 \ \vec{0} \ \vec{0}] P^{-1} = \begin{bmatrix} 7 & 0 & 0 \\ 5 & 0 & 0 \\ 8 & 0 & 0 \end{bmatrix} \frac{1}{60} \begin{bmatrix} 3 & 3 & 3 \\ 44 & -36 & -16 \\ 5 & -15 & 5 \end{bmatrix}$$

$$= \frac{1}{20} \begin{bmatrix} 7 & 7 & 7 \\ 5 & 5 & 5 \\ 8 & 8 & 8 \end{bmatrix}$$

Example 4. Find $\lim_{t\to\infty} A^t \vec{x}_0$

$$\lim_{t\to\infty} A^t \vec{x}_0 = \frac{1}{20} \vec{b}_1 = \frac{1}{20} \begin{bmatrix} 7 \\ 5 \\ 8 \end{bmatrix} = \begin{bmatrix} 7/20 \\ 5/20 \\ 8/20 \end{bmatrix}$$

Let's summarize the example:

Equilibria for regular transition matrices:

Definition.

1. A vector $\vec{x} \in \mathbb{R}^n$ is said to be a **distribution** vector if its entries are non-negative and the sum is 1.
2. A square matrix A is said to be a **transition matrix** (or stochastic matrix) if all its columns are distributions vectors. The corresponding dynamical system is also called **Markov Process**.
3. A transition matrix is said to be **regular** (or eventually positive) if the matrix A^m is positive for some integer $m > 0$.

Theorem.

Let A be a regular transition $n \times n$ matrix.
1. There exists **exactly** one **distribution** vector $\vec{x} \in \mathbb{R}^n$ such that

$$A\vec{x} = \vec{x}$$

which is called **equilibrium** distribution for A denoted as $\vec{x}_{equ}$.
2. If $\vec{x}_0$ is any distribution vector in $\mathbb{R}^n$, then

$$\lim_{m\to\infty} (A^m \vec{x}_0) = \vec{x}_{equ}$$

3. The columns of $\lim_{m\to\infty}(A^m)$ are all $\vec{x}_{equ}$, that is

$$\lim_{m\to\infty} (A^m) = [\vec{x}_{equ}\ \vec{x}_{equ} \dots \vec{x}_{equ}]$$

Remark: So, $\vec{x}_{equ}$ is the **distribution** eigenvector with eigenvalue 1.

Now, let us use the theorem to redo example 3 and example 4.

Redo Example 4: For any distribution vector $\vec{x}_0$, find $\lim\limits_{t\to\infty} A^t \vec{x}_0$

$$A - I = \begin{bmatrix} -0.3 & 0.1 & 0.2 \\ 0.2 & -0.6 & 0.2 \\ -0.1 & 0.5 & -0.4 \end{bmatrix} \to \cdot \to \textbf{rref}$$

So a basis for $E_{\lambda=1}$ is $\{ \begin{bmatrix} 7 \\ 5 \\ 8 \end{bmatrix} \}$ This eigenvector vector is not a distribution vector.

So the distribution eigenvector is $\vec{x}_{equ} = \dfrac{1}{7+5+8} \begin{bmatrix} 7 \\ 5 \\ 8 \end{bmatrix} = \begin{bmatrix} 7/20 \\ 5/20 \\ 8/20 \end{bmatrix}$

So, for any distribution vector $\vec{x}_0$, $\lim\limits_{t\to\infty} A^t \vec{x}_0 = \vec{x}_{equ} = \begin{bmatrix} 7/20 \\ 5/20 \\ 8/20 \end{bmatrix}$

Redo Example 3: Find $\lim\limits_{t\to\infty} A^t$

By theorem,

$$\lim_{t\to\infty} A^t = [\vec{x}_{equ}\ \vec{x}_{equ}\ \vec{x}_{equ}] = \begin{bmatrix} 7/20 & 7/20 & 7/20 \\ 5/20 & 5/20 & 5/20 \\ 8/20 & 8/20 & 8/20 \end{bmatrix}$$

Further reading about the PageRank:

Other lectures:

http://pi.math.cornell.edu/ mec/Winter2009/RalucaRemus/Lecture3/lecture3.html

A little more professional:

https://www.rose-hulman.edu/ bryan/googleFinalVersionFixed.pdf

https://www.math.purdue.edu/ eremenko/dvi/google-eigenvector.pdf

http://www.ams.org/publicoutreach/feature-column/fcarc-pagerank

Original paper:

Sergey Brin and Lawrence Page http://infolab.stanford.edu/ backrub/google.html

Example 5. Application in weather predicting We want to predict the weather in Boston (baby level). Consulting local weather records over the past decade, we find that

(a) If today is sunny, there is a 70% chance that tomorrow will also be sunny, 20% chance cloudy, 10% rainy.

(b) If today is cloudy, the chances are 80% that tomorrow will also be cloudy, 10% chance sunny and 10% chance rainy.

(c) If today is rainy, the chances are 50% that tomorrow will also be rain, 30% cloudy and 20% sunny.

Question: given that today is sunny, what is the probability that next Saturdays weather will also be sunny?

$$B = \begin{bmatrix} 0.7 & 0.8 & 0.5 \\ 0.2 & 0.1 & 0.3 \\ 0.1 & 0.1 & 0.2 \end{bmatrix}$$

Eigenvalues of B are 1, 0.1, -0.1.

The equilibrium vector is $\begin{bmatrix} 23/33 \\ 19/99 \\ 1/9 \end{bmatrix} \approx \begin{bmatrix} 0.697 \\ 0.192 \\ 0.111 \end{bmatrix}$

Example 6. There is a bicycle sharing company in MA. Records indicate that, on average, 10% of the customers taking a bicycle in downtown go to Cambridge and 30% go to suburbs. Customers boarding in Cambridge have a 30% chance of going to downtown and a 30% chance of going to the suburbs, while suburban customers choose downtown 40% of the time and Cambridge 30% of the time. The owner of the bicycle sharing company is interested in knowing where the bicycle will end up, on average.

The transition matrix is $A = \begin{bmatrix} 0.6 & 0.3 & 0.4 \\ 0.1 & 0.4 & 0.3 \\ 0.3 & 0.3 & 0.3 \end{bmatrix}$ Note that A is regular. The equilibrium eigenvector is $\begin{bmatrix} 0.4714 \\ 0.2286 \\ 0.3 \end{bmatrix}$