PDEs and Graph Based Learning

Summer School on Random Structures in Optimizations and Related Applications

Lecture 2: PageRank

Instructor: Jeff Calder (jcalder@umn.edu)

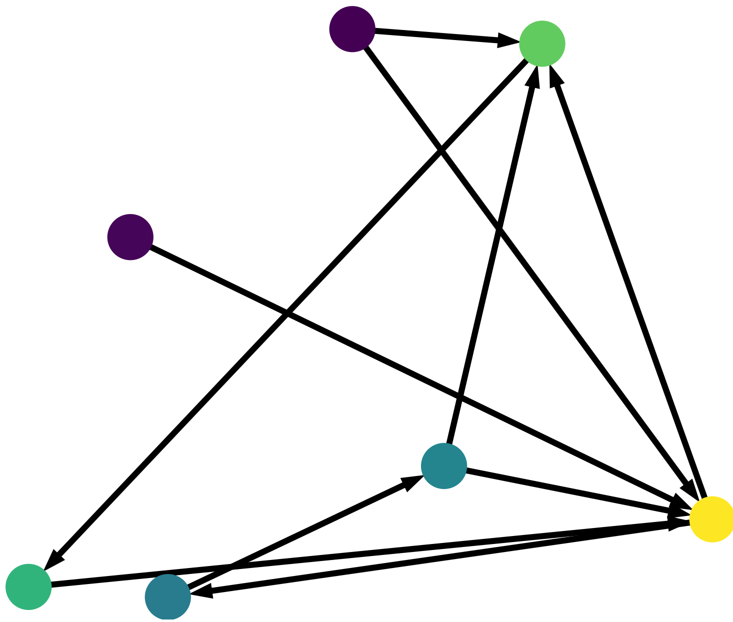
Web: http://www-users.math.umn.edu/~jwcalder

Lecture Notes: http://www-users.math.umn.edu/~jwcalder/5467S21

PageRank

The PageRank algorithm ranks websites based on the link structure of the internet. It was used to sort Google search results until 2006, and has been used in

- Biology (GeneRank), chemistry, ecology, neuroscience, physics, sports, and computer systems. . .



PageRank

Main Idea: Take a random walk on the internet for T steps.

$$\text{Rank of site } i = \lim_{T\to\infty} \frac{1}{T}(\text{Number of times site } i \text{ is visited}).$$

Problem: Random walks can get stuck in disconnected components of the internet, and may never visit a given site i.

Solution: Every so often, the random walker teleports to a random site on the internet. The walker is called a random surfer.

Code demo

Mathematics of PageRank

To describe PageRank mathematically, we start with an adjacency matrix W

$$W(i,j) = \begin{cases} 1, & \text{if site } i \text{ links to site } j \\ 0, & \text{otherwise.} \end{cases}$$

We also have a probability transition matrix P for the random walk:

$$P(i,j) = \text{Probability of stepping from } j \text{ to } i.$$

Both P and W are $n \times n$ matrices, $n =$number of webpages.

Mathematics of PageRank

Clicking on a link at random from webpage j leads to the transition probabilities

$$P(i,j) = \frac{W(j,i)}{\sum_{k=1}^{n} W(j,k)}.$$

Exercise 1. Show that $P = W^T D^{-1}$, where D is the diagonal matrix with diagonal entries $D(i,i) = \sum_{j=1}^{n} W(i,j)$. △

Random surfer

Let $\alpha \in [0, 1)$ be the random walk probability, and let $v \in \mathbb{R}^n$ be the teleportation probability distribution. That is, $v(i) \geq 0$ for all i, and $\sum_i v(i) = 1$.

Random surfer dynamics: When at website j, the random surfer chooses the next site as follows:

1. With probability α the surfer clicks an outgoing link at random, that is, the surfer navigates to website i with probability $P(i, j)$.

2. With probability $1 - \alpha$ the surfer teleports to website i with probability $v(i)$.

Teleportation

Teleportation distribution: Common choices are

- $v(i) = 1/n$ for all i (jump to a site uniformly at random).
- (Localized PageRank) $v(i) = \delta_{ij}$ (always jump back to site j).

Localized PageRank ranks all sites based on their similarity to site j.

The PageRank vector

For $k \geq 0$ define

$$x_k(i) = \text{Probability that the random surfer is at page } i \text{ on step } k.$$

Definition 2. The **PageRank** vector x is

$$x(i) = \lim_{k \to \infty} x_k(i),$$

provided the limit exists.

Transition probabilities

To see how x_k transitions to x_{k+1} requires some probability. We condition on the location of the surfer at step k, and on the outcome of the coin flip, to obtain

$$x_{k+1}(i) = (1-\alpha)v(i) + \alpha \sum_{j=1}^{n} P(i,j) x_k(j).$$

We can write this in matrix/vector form as

$$x_{k+1} = (1-\alpha)v + \alpha P x_k. \tag{1}$$

If x_k converges to a vector x as $k \to \infty$, then x should satisfy

$$x = (1-\alpha)v + \alpha P x.$$

Question: Does x_k converge as $k \to \infty$, and if so, how quickly does it converge?

Analysis of PageRank

We consider the PageRank equation

$$x = (1-\alpha)v + \alpha P x. \tag{2}$$

Lemma 3. *Let $v \in \mathbb{R}^n$ and $0 \leq \alpha < 1$. Then there is a unique vector $x \in \mathbb{R}^n$ solving the PageRank equation (2). Furthermore, the following hold.*

(i) We have $\sum_{i=1}^n x(i) = \sum_{i=1}^n v(i)$.

(ii) If $v(i) \geq 0$ for all i, then $x(i) \geq 0$ for all i.

The ℓ_1-norm

It will be more convenient to work in the ℓ_1-norm $\|\cdot\|_1$ defined by

$$\|x\|_1 = \sum_{i=1}^{n} |x(i)|.$$

In the ℓ_1-norm, the transition matrix P is non-expansive.

Proposition 4. *We have* $\|Px\|_1 \leq \|x\|_1$.

Proof:
$$\|Px\|_1 = \sum_{i=1}^{n} |Px(i)|$$
$$= \sum_{i=1}^{n} \left| \sum_{j=1}^{n} P(i,j)\, x(j) \right|$$
$$\leq \sum_{i=1}^{n} \sum_{j=1}^{n} P(i,j)\, |x(j)|$$
$$= \sum_{j=1}^{n} |x(j)| \sum_{i=1}^{n} P(i,j)$$

$$\underbrace{\qquad\qquad}_{=1}$$

$$= \|x\|_1 \qquad \blacksquare$$

Proof (of lemma 3) $\quad x = (1-\alpha)v + \alpha P x$

$$x - \alpha P x = (1-\alpha) v$$

$$\underbrace{(1-\alpha)^{-1}(I - \alpha P)}_{A} x = v$$

$$Ax = v$$

Let $z \in \mathrm{Ker}(A)$, so $Az = 0$ or $z = \alpha P z$

$$\|z\|_1 = \|\alpha P z\|_1 = \alpha \|Pz\|_1$$

$$\underset{\text{prop 4}}{\leq} \alpha \|z\|_1$$

$$(1-\alpha)\|z\|_1 = 0 \qquad \text{since } 1-\alpha \geq 0$$

Provided $0 \leq \alpha < 1$ we have $z = 0$

and $\operatorname{Ker}(A) = \{0\}$.

Thus, for every v, $\exists!\, x \in \mathbb{R}^n$ solving the PageRank problem.

(i) $$\sum_{i=1}^{n} x(i) = \sum_{i=1}^{n} \left((1-\alpha) v(i) + \alpha \sum_{j=1}^{n} P(i,j)\, x(j) \right)$$

$$= (1-\alpha) \sum_{i=1}^{n} v(i) + \alpha \sum_{j=1}^{n} x(j) \underbrace{\sum_{i=1}^{n} P(i,j)}_{=1}$$

$$= (1-\alpha) \sum_{i=1}^{n} v(i) + \alpha \sum_{i=1}^{n} x(i)$$

$$(1-\alpha) \sum_{i=1}^{n} x(i) = (1-\alpha) \sum_{i=1}^{n} v(i)$$

Provided $\alpha < 1$, $\sum x(i) = \sum v(i)$.

(ii) $$|x(i)| = \left| (1-\alpha) v(i) + \alpha \sum_{j=1}^{n} P(i,j) x(j) \right|$$

$v(i) \geq 0$ $$\leq (1-\alpha) v(i) + \alpha \sum_{j=1}^{n} P(i,j) |x(j)|$$

Sum over i on both sides

$$\sum_{i=1}^{n} |x(i)| \le (1-\alpha)\sum_{i=1}^{n} v(i) + \alpha \sum_{j=1}^{n} |x(j)|$$

$$(1-\alpha)\sum_{i=1}^{n} |x(i)| \le (1-\alpha)\sum_{i=1}^{n} v(i)$$

$\alpha < 1$ by ii) $= (1-\alpha)\sum_{i=1}^{n} x(i)$

$$\sum_{i=1}^{n} \underbrace{\left(|x(i)| - x(i)\right)}_{\ge 0} \le 0$$

$$\Rightarrow \quad |x(i)| = x(i) \Rightarrow x(i) \ge 0$$

$\blacksquare$

Eigenvector problem

(*) $x = (1-\alpha) v + \alpha P x$

Remark 5. When v is a probability distribution, it is common to re-write the PageRank problem (2) as an eigenvector problem

$$P_\alpha x = x$$

where

$$P_\alpha := (1-\alpha) v \mathbf{1}^T + \alpha P.$$

(*) Assume $x(i) \geq 0$, $\mathbf{1}^T x = 1$ (prob. dist.)

$\Rightarrow x = (1-\alpha) v \mathbf{1}^T x + \alpha P x$

$$= \underbrace{\left[(1-\alpha) v \mathbf{1}^T + \alpha P \right]}_{= P_\alpha} x$$

Convergence of the PageRank iteration

Let $v \in \mathbb{R}^n$ and $0 \leq \alpha < 1$. Let x_k satisfy the PageRank iteration

$$x_{k+1} = (1 - \alpha)v + \alpha P x_k,$$

and let x be the unique solution of the PageRank problem

$$x = (1 - \alpha)v + \alpha P x.$$

Theorem 6. *We have*

$$\|x_k - x\|_1 \leq \alpha^k \|x_0 - x\|_1. \tag{3}$$

Since $0 \leq \alpha < 1$, this is convergence of $x_k \to x$ with a linear convergence rate of α.

Proof:

$$x_k = (1-\alpha)v + \alpha P x_{k-1}$$
$$x = (1-\alpha)v + \alpha P x$$

Sub. equations

$$x_k - x = \alpha P x_{k-1} - \alpha P x$$

$$= \alpha P (x_{k-1} - x)$$

$$\| x_k - x \| = \alpha \| P (x_{k-1} - x) \|$$

$$\|Px\|_1 \le \|x\|_1$$

$$\le \alpha \| x_{k-1} - x \|$$

$$\le \alpha^2 \| x_{k-2} - x \|$$

$$\le \alpha^3 \| x_{k-3} - x \|$$

$$\vdots$$

$$\leq \alpha^k \| x_0 - x \| \qquad \blacksquare$$

Graph Convolutional Neural Networks
(GCN) (Kipf / Welling 2016)

$$X^{k+1} = \sigma(P X^k W^k)$$

$$\sigma(t) = \text{ReLU}(t)$$

$$= \max(0, t)$$

$$\sigma(t) = \frac{1}{1+e^{-t}}$$

Power iteration

Remark 7. In the eigenvector formulation discussed above, the PageRank iteration $x_{k+1} = P_\alpha x_k$ is basically the power iteration to find the largest eigenvector of P. The normalization step is not needed since $\|x_k\|_1 = 1$ for all k.

$$x_{k+1} = P_\alpha x_k$$

.... $x_k = P_\alpha^k x_0 \longrightarrow$ largest eigenvector

Personalized PageRank for image retrieval (.ipynb)