Personal Pageranks & Spilling Paint

15-859N
12/6/16

$G=(V,E)$ undir (weighted)

$A \equiv$ adj $\quad D \equiv$ deg matrix

<u>Def</u> $S \subseteq V \quad vol(S) = d(S) = \sum_{v \in S} d(v)$

$$\partial S = \{(v,w) \mid v \in S \ \& \ w \notin S\}$$

$$\text{Conductance}(S) = \frac{|\partial S|}{d(S)}$$

Goal: Given $v \in V$ find small $v \in S$ with small conduct.

Idea: Use "local" random walks to find S.

Goal: Local page rank (Google)

Two random walks

Regular or eager

$$p^{(t+1)} = A^T D^{-1} p^{(t)} \qquad W \equiv A^T D^{-1}$$
$$= W p^{(t)}$$

Lazy

$$p^{(t+1)} = \frac{1}{2}(I + W)\, p^{(t)}$$
$$\equiv \hat{W} p^{(t)}$$

Walks with resets

Let u be a distribution over V

i.e. $0 \le u \in \mathbb{R}^n$ $|u|_1 = 1$

eg u is characteristic vector of $v \in V$

$$p^{(t+1)} = \alpha u + (1-\alpha) W p^{(t)}$$

or

$$= \alpha u + (1-\alpha) \hat{W} p^{(t)}$$

$$0 < \alpha < 1$$

<u>Thm</u> $\exists!$ $P_u = \alpha u + (1-\alpha) W P_u$

<u>pf</u> rewrite

$$P_u - (1-\alpha) W P_u = \alpha u$$

$$\left[I - (1-\alpha) W\right] P_u = \alpha u$$

Claim $(I - (1-\alpha) W)$ is non-sing

$$|\lambda(W)| \leq 1$$

$$\Rightarrow |\lambda((1-\alpha) W)| \leq 1-\alpha < 1$$

$$\Rightarrow |\lambda(I - (1-\alpha) W)| > 0$$

thus $P_u = (I - (1-\alpha) W)^{-1} \alpha u$

Recall Sym $X \in \mathbb{R}^{n \times n}$ $|\lambda(X)| < 1$

then $(I - X)^{-1} = (I + X + X^2 + X^3 + \cdots)$

thus $P_u = \alpha \sum_{t \geq 0} (1-\alpha)^t W^t u$

Spilling Paint View

Idea: bucket of paint at vertex v.
say remaining paint $r^{(t)}$.
we spill $\alpha r^{(t)}$ it dries/sticks

thus our evolution is:
let $s^{(t)}, r^{(t)}$ be dist of paint.

$$s^{(t+1)} = s^{(t)} + \alpha r^{(t)}$$
$$r^{(t+1)} = (1-\alpha) W r^{(t)}$$

Where is the stuck paint?

$$s^{\infty} = \alpha \sum_{t \geq 0} r^{(t)} = \alpha \sum_{t \geq 0} (1-\alpha)^t W^t r^{(0)}$$
$$= \alpha \sum_{t \geq 0} (1-\alpha)^t W^t \mu$$

---

thus $P_\mu = s^{\infty}$!

<u>Claim</u> Reg & Lazy Walk diff by a constant.

$$s^{\infty} = \alpha\left(I - (1-\alpha)\hat{W}\right)^{-1} u \qquad \hat{W} = I/2 + W/2$$

$$= \alpha\left(\frac{1+\alpha}{2} I - \left(\frac{1-\alpha}{2}\right) W\right)^{-1} u$$

$$= \frac{2\alpha}{1+\alpha}\left(I - \left(\frac{1-\alpha}{1+\alpha}\right) W\right)^{-1} u$$

$$\beta = \frac{2\alpha}{1+\alpha} \quad \text{note} \quad 1-\beta = 1 - \frac{2\alpha}{1+\alpha} = \frac{1-\alpha}{1+\alpha}$$

$$= \beta\left(I - \beta W\right)^{-1} u$$

Pushing wet paint

Suppose s dry paint
r wet

$$p_{s,r} = s + \alpha \sum_{t \geq 0} (1-\alpha)^t W^t r = s + \alpha (I - (1-\alpha) W)^{-1} r$$

Consider a partial update at some u.

Update(u)

$$s'(u) = s(u) + \alpha r(u)$$

$$p'(u) = 0$$

$$p'(v) = p(v) + \frac{1-\alpha}{d(u)} p(u) \quad \forall v \in N(u)$$

Lemma $P_{s',r'} = P_{s,r}$

pf see notes

Alg Approx Painting

Init: $s \equiv 0$ $r = \chi_u$

While $\exists$ $r(u) \geq \varepsilon d(u)$

Pick max $r(u)/d(u)$

Update(u)

Computing small conductance sets

$S \subseteq V$

Def $\mathrm{Vol}(S) = d(S) = \sum_{v \in S} d(v)$

$$\partial S = \{(x,y) \mid x \in S \ \& \ y \notin S\}$$

$$|\partial S| = \sum_{e \in \partial S} w(e)$$

Conductance $\Phi(S) = \dfrac{|\partial(S)|}{\min\{d(S), d(\bar{S})\}} = \phi(S)$

$$\chi_S(u) \overset{\text{def}}{=} \begin{cases} 1 & \text{if } u \in S \\ 0 & \text{o.w.} \end{cases}$$

Let P be our stationary dist with resets starting at v.

$$q_v(u) = \frac{P_v(u)}{d(u)}$$

sort $q_v(u)$ giving $q_v(1) \geq q_v(2) \geq \cdots \geq q_v(n)$

<u>Def</u> $S_k = \{1, \ldots, k\}$ $\quad S_k \subseteq V$

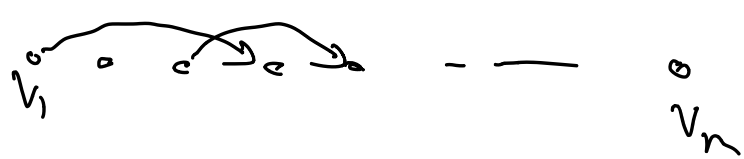
Lemma 5.1 $\displaystyle\sum_{\substack{i \leq k < j \\ (i,j) \in E}} q(i) - q(j) \leq \alpha \qquad S \subseteq S_k$

The stationary as a flow (circulation)

Note $g(i) = \dfrac{p(i)}{d(i)} \cong$ prob of leaving v_i

$\underset{(i,j)\in E}{(1-\alpha)\big[g(i)-g(j)\big]}$ can be view as a circulation f_{ij} when we include the resets.

Since $g(i) \geq g(j)$ for $i \leq j$ all flow on graph is left to right



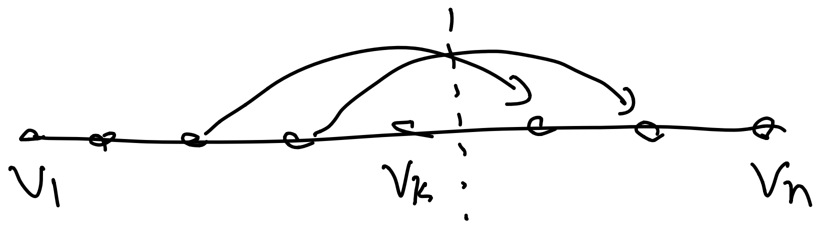
the right to left is reset flow.

Claim: V_1 is reset vertex u.

<u>pf</u> All graph flow is out of V_1

Thus V_1 must be the reset vertex.

Consider graph flow crossing V_k



$$\text{flow} = \text{reset for } V_{k+1} \text{ to } V_n$$

$$= \alpha(P_{k+1} + \cdots P_n) \quad (*)$$

Observe that $P_1 \geq \alpha$

$$(*) \leq \alpha(1-\alpha)$$

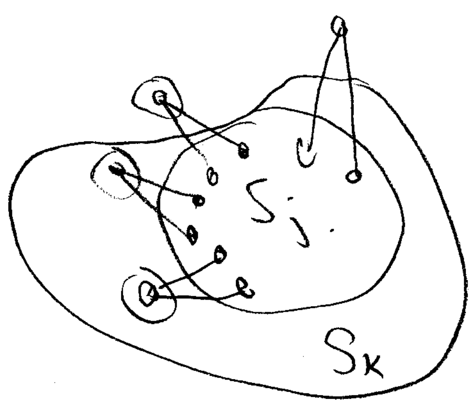
This proves Lemma 5.1

Lemma 5.2 $\phi(S_j) \geq 2\theta$ Then $\exists k > j$ s.t

1) $d(S_k) \geq (1+\theta)\, d(S_j)$

2) $g(j) - g(k) \leq \dfrac{\alpha}{\theta d(S_j)}$

Pf Set $k = \operatorname{argmin}_k \left\{ |\partial S_j \cap E(S_k)| \geq \dfrac{|\partial S_j|}{2} \right\}$



$|\partial(S_j)| \geq 2\theta\, d(S_j)$

ok

$$\alpha \geq \sum_{ab \in E;\, a \leq j;\, b \geq k} g(a) - g(b) \geq \theta d(S_j)(g(j) - g(k))$$

This gives 2)

5.3

Lemma $\phi(S_j) \geq 2\theta \ \forall j$ st $d(S_j) \leq 2m/3$

$$h = \arg\min_h \{ d(S_h) \geq 2m/3 \}$$

then $\forall i \leq h$ $\quad g(h) \geq g(i) - \dfrac{2\alpha}{\theta^2 d(S_i)}$

Let $K_1, \ldots K_t$ be seq of K from last lemma

$$g(i) - g(K_1) \leq \frac{\alpha}{\theta d(S_i)}$$

$$g(K_1) - g(K_2) \leq \frac{\alpha}{\theta d(S_{K_1})} \leq \frac{\alpha}{\theta(1+\theta) d(S_i)}$$

$$\vdots$$

$$g(i) - g(h) \leq \frac{\alpha}{\theta d(S_i)} \left(1 + \frac{1}{1+\theta} + \frac{1}{(1+\theta)^2} + \cdots \right)$$

$$\leq \frac{\alpha}{\theta d(S_i)} \frac{1+\theta}{\theta} \qquad \theta \leq 1$$

$$\leq \frac{2\alpha}{\theta^2 d(S_i)}$$

$S \subseteq V$

$$\pi_S(u) = \begin{cases} \frac{d(u)}{d(S)} & \text{if } u \in S \\ 0 & \text{o.w.} \end{cases}$$

Note: $\rho_{\pi_S} = \alpha \sum (1-\alpha)^t W^t \pi_S$

Claim $\chi_{\bar{S}}^T W^t \pi_S \leq t\,\phi(S)$

case $t=1$

$$\sum_{v \notin S} \sum_{uv \in \partial S} \frac{1}{d(u)} \cdot \frac{d(u)}{d(S)} = \frac{1}{d(S)} \sum_{v \notin S} \sum_{uv \in \partial S} 1 = \frac{|\partial S|}{d(S)} = \phi(S)$$

to check $t > 1$

Lemma $\chi_{\bar{S}}^T \, p_{\pi_S} \leq \phi(S) \, \frac{1-\alpha}{\alpha}$

$$\text{RHS} = \alpha \sum_{t \geq 0} (1-\alpha)^t \, \chi_{\bar{S}}^T W^t \pi_S$$

$$\leq \alpha \sum_{t \geq 0} (1-\alpha)^t \, t \, \phi(S)$$

$$= \alpha \, \phi(S) \sum_{t \geq 0} (1-\alpha)^t \, t \qquad (*)$$

$\beta = 1-\alpha$

$$\frac{1}{(1-\beta)} = \sum_{t \geq 0} \beta^t \quad \Rightarrow \quad + \frac{1}{(1-\beta)^2} = \sum_{t \geq 1} \beta^{t-1} \cdot t$$

$$\frac{\beta}{(1-\beta)^2} = \sum_{t \geq 1} \beta^t \cdot t = \sum_{t \geq 0} \beta^t t$$

$1-\beta = \alpha$

$$(*) = \alpha \, \phi(S) \left(\frac{1-\alpha}{\alpha^2} \right) = \phi(S) \left(\frac{1-\alpha}{\alpha} \right)$$

Set α st $\phi(S)\left(\frac{1-\alpha}{\alpha}\right) < 1/3$

<u>Lemma</u> $\exists i$ st $d(S_i) \leq d(S)$ and

$$g(i) \geq \frac{2/3}{d(S_i)\, H_{2m}}$$

$\gamma = H_{2m}$

<u>pf.</u> ?

Thm

Suppose the following for some S

1) $d(S) \le 2m/9\gamma \qquad \gamma = H(2m)$

2) start walk with resets for $u \in S$

   prob $= \pi_S$

3) Set α st $\phi(S)\left(\frac{1-\alpha}{\alpha}\right) < 1/3 \approx \phi(S) < \frac{\alpha}{3}$

then $\exists j$ st $\phi(S_j) = O\left(\sqrt{\phi(S) \log m}\right)$ (*)

pf Suppose false

set $\sqrt{6\alpha\gamma} \approx \sqrt{2\phi(S) \ln n} = \Theta$

$\forall j \quad \phi(S_j) \ge 2\Theta$

$g(h) \ge g(i) - \dfrac{2\alpha}{\Theta^2 d(S_i)} = g(i) - \dfrac{2\alpha}{6\alpha\gamma\, d(S_i)} \ge$

$$\frac{2/3}{d(S_i)\gamma} - \frac{1/3}{\gamma d(S_i)} = \frac{1/3}{\gamma d(S_i)}$$

$\forall i \le h \quad g(i) \ge g(h) \ge \dfrac{1}{3\gamma d(S_i)}$

$$\sum_{i=1}^{n} d(i)\, g(i) = 1$$

$$\sum_{i=1}^{h} d(i)\, g(i) \ge g(h) \sum_{i=1}^{h} d(i) = d(S_h)\, g(h)$$

$$\ge \left(\tfrac{2}{3}m\right) \Big/ 3\gamma d(S_i) = \frac{2m}{9\gamma d(S_i)} \ge \frac{2m}{9\gamma d(S)} = (*)$$

$d(S) < \dfrac{2m}{9\gamma} \Rightarrow (*) > 1$ contra!