15-750
4/19/19

Personal Pageranks & Spilling Paint

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

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

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

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

$$\text{Conductance}(S) = \frac{|\partial S|}{\min\{d(S), d(V \setminus S)\}}$$

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 Walk.

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

Lazy Walk.

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

Walks with resets

Let u be a distribution over V

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

eg u is characteristic vector of $v \in V$, χ_v.

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

or

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

$$0 < \alpha < 1$$

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

pf rewrite

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

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

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

We show all eigenvalues of $(I - (1-\alpha) W) > 0$

Let $M \in \mathbb{R}^{n \times n}$

Def: $\lambda(M)$ = set of all eigenvalues of M.

$$|\lambda(M)| = \max\{|\lambda| \mid \lambda \in \lambda(M)\}$$

Note In general eigenvalues & vectors may be complex.

Def $|a+ib| = \sqrt{a^2+b^2}$

<u>Lemma:</u> $M \in \mathbb{R}^{n \times n}$, $M \geq 0$,
s = max row sum then $|\lambda(M)| \leq s$.

Pf Suppose $\lambda \in \lambda(M)$ i.e. $MX = \lambda X$ some X.
WLOG $|X_n| \geq \{|X_1|, \cdots, |X_{n-1}|\}$
Let S_n = sum of row n in M.
Now $\left(\frac{1}{S_n} MX\right)_n = \lambda / S_n X_n$
LHS is a convex combination of $\{X_1, \cdots, X_n\}$.
Thus $|\lambda / S_n| \leq 1$ or $|\lambda| \leq S_n \leq S$ $\square$

Note: $W = AD^{-1} \geq 0$ & each row sum $= 1$.
Thus: $\forall \lambda$ $|\lambda(W)| \leq 1$
$\Rightarrow |\lambda((1-\alpha)W)| \leq 1-\alpha < 1$
$\Rightarrow |\lambda(I-(1-\alpha)W)| > 0$ (*)

<u>Thus:</u> $P_u = (I-(1-\alpha)W)^{-1} \alpha u$ since inverse exists

---

To see (*) note $MX = \lambda X$ then
$(I+M)X = X + MX = X + \lambda X = (1+\lambda)X$

Spectral Thm

Recall the spectral thm.

Thm Sym $M \in \mathbb{R}^{n \times n}$ then $M = U^T \Lambda U$ where $U^T U = I$ & Λ is diagonal.

Thus if $M \in \mathbb{R}^{n \times n}$ & $|\lambda(M)| < 1$

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

Note $W = AD^{-1}$ is not sym but still

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

Thus P_u can be expressed as a sum of deprecated (lossy) random walks from u.

Spilling Paint View

Idea: 1) Start with a bucket of paint at vertex v.

2) Let $r^{(t)}$ be wet paint at time t.

3) At time t we spill $\alpha r^{(t)}$ paint it dries/sticks

4) We move the remaining wet paint to a random neighbor.

Let $s^{(t)}, r^{(t)}$ be dist of dry/wet paint.
Thus our evolution equations are:

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

$$r^{(t+1)} = (1-\alpha) W r^{(t)} \quad \text{or} \quad r^{(t)} = (1-\alpha)^t W^t r^{(0)}$$

Where is the stuck paint?

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

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

Thus $P_n = S^{\infty}$!

Claim: Reg & Lazy walks differ by at most a constant for local page rank.

Recall Eager $W = A^T D^{-1}$

Lazy $\hat{W} = I/2 + W/2$

Final stuck paint

$$S^{\infty}(\text{Lazy}, \alpha) = \alpha\left(I - (1-\alpha)\hat{W}\right)^{-1} u$$

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

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

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

Set $\beta = \dfrac{2\alpha}{1+\alpha}$ then $1-\beta = 1 - \dfrac{2\alpha}{1+\alpha} = \dfrac{1-\alpha}{1+\alpha}$

Thus $$= \beta\left(I - (1-\beta)W\right)^{-1} u = S^{\infty}(\text{Eager}, \beta)$$

Pushing wet Paint

Suppose we have s dry & r wet paint.

If we run Paint-Spilling from (s,r) we get.

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

$$= s + \alpha (I - (1-\alpha)W)^{-1} r \quad \text{(dry paint)}$$

Consider a partial update at some $u \in V$.

<u>Def</u> $\mathrm{Update}(v) \equiv$ (unweighted case)

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

$$r'(v) = 0$$

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

Here we are doing eager walks.

Let $(s', r') = \mathrm{Update}_u(s, r)$

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

pf See Spielman's notes.

Alg: Approx Painting(u, α, ε)

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

While: $\exists v \in V$ $r(v) \geq \varepsilon\, d(v)$

Pick $\operatorname{argmax}_{v \in V} r(v)/d(v)$

Update(v)

Thm Alg ApproxPainting terminates after $1/\varepsilon\alpha$ iterations.

(No pf)

Understanding the stationary P_v for walks with resets from v.

Consider $q_u(v) = P_u(v)/d(v)$

$d(v) \equiv$ degree of v.

Sort q_u giving $q_1(v) \geq q_2(v) \geq \cdots \geq q_n(v)$

Let S_k = vertices $\{1, \cdots, k\}$

We consider two types of flow in the stationary P_u.

1) Reset-flow: flow back to u.
2) Graph-flow: flow on edges of G.

Note Total flow is a circulation!

The stationary as a flow (circulation)

<u>Note</u> $g(i) = \dfrac{p(i)}{d(i)} \cong$ prob of leaving V_i

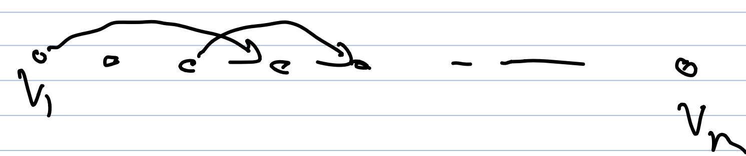
on some fixed edge out of V_i.

$$f_{ij} = (1-\alpha)\left[g(i) - g(j)\right] \cong \text{net prob of traversing}$$
$$(i,j) \in E \qquad\qquad (i,j) \in E.$$

Let f_{ij} be the graph-flow.

Since $g(i) \geq g(j)$ for $i \leq j$ all

Graph-flow on edges is left to right.



Since graph-flow + reset-flow is a circulation.

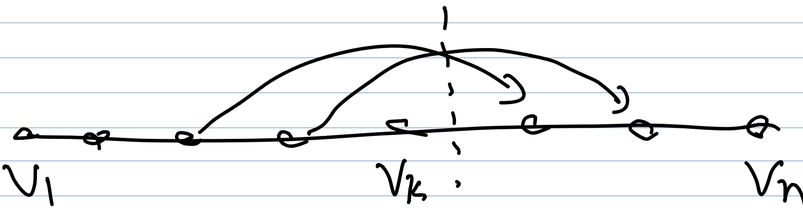
Intuitively the reset-flow should be from right to left.

<u>Claim:</u> V_1 is the 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 $P_1 \geq \alpha$

Main Thm

In words: If $C \subseteq V$ is of low conductance and we do LPR from random $u \in C$ then one of the sets S_k will have low conductance.

Thm $C \subseteq V$, $d(C) \leq d(V)/2$, $\Phi(C) \leq \phi$

$\alpha = \phi / c \cdot \log n$. If we pick u by degree in C then with prob $\geq 1/2$

$\exists k$ st

1) $\Phi(S_k) \leq O(\sqrt{\phi(C) \log m})$

2) $2/3\, d(S_k) \leq d(S_k \cap C)$