15-859N
12/2/11

Local Graph Partitioning using PageRank Vectors

Prelims

$G = (V, E)$ undir & unweighted

D = deg matrix & A adj

Def Stationary dist $S \subseteq V$

$$\psi_S(x) = \begin{cases} \frac{d(x)}{vol(S)} & x \in S \\ 0 & \text{o.w.} \end{cases}$$

indicator fcn $\chi_v(x) = \begin{cases} 1 & \text{if } x = v \\ 0 & \text{o.w.} \end{cases}$

$p^{1 \times n}$ vector

Support $Supp(p) = \{v \in V \mid p(v) \neq 0\}$

$$p(S) = \sum_{u \in S} p(u)$$

Random walks with resets

Standard transition matrix $D^{-1}A$

$d \equiv 1\times n$ vector trans $d' = dD^{-1}A$

Lazy random walk

$$W = \frac{1}{2}(I + D^{-1}A)$$

Def PageRank vector $s \equiv 1\times n$ vector

$$pr_\alpha(s) = \alpha s + (1-\alpha)\, pr_\alpha(s) W$$

(Brin & Page) used $s = \vec{1}/n$

Here consider $s = \chi_v$

Prop 1 $\forall s, \alpha$ $0 < \alpha \le 1$ $\exists!\, pr_\alpha(s)$

satisfying $pr_\alpha(s) = \alpha s + (1-\alpha)\, pr_\alpha(s) W$ $\quad (*)$

pf

solutions to

$$p\underbrace{\left(I-(1-\alpha)W\right)}_{M} = \alpha S$$

$$M = I - \frac{(1-\alpha)I}{2} - \frac{(1-\alpha)D^{-1}A}{2}$$

$$= \underbrace{\left(\frac{1+\alpha}{2}\right)I - \underbrace{\left(\frac{1-\alpha}{2}\right)\underbrace{D^{-1}A}_{\text{col sums}=1}}_{\text{col sums}=\frac{1-\alpha}{2}}}_{\text{strictly diagonally Dom}}$$

M is non singular!

Note $S=0$ then $pr_\alpha(s)=0$

Note S need not be nonnegative!

$$p = \alpha S M^{-1}$$

Prop 2 consider

$$R_\alpha = \alpha I + \alpha \sum_{t=1}^{\infty} (1-\alpha)^t W^t$$

then $pr_\alpha(s) = sR_\alpha$

toshow: sR_α sat $(*)$

$$= \alpha s + (1-\alpha)\, sR_\alpha W$$

$$= \alpha s + (1-\alpha)\, s\left(\alpha I + \alpha \sum_{t=1}^{\infty} (1-\alpha)^t W^t\right) W$$

$$= \alpha s + s\alpha\left((1-\alpha) W + \sum_{t=2}^{\infty} (1-\alpha)^t W^t\right)$$

$$= \alpha s + s\alpha\left(\sum_{t=1}^{\infty} (1-\alpha)^t W^t\right)$$

$$= s R_\alpha$$

<u>Cor</u> $pr_s(s)$ linear in s.

$$Vol(S) = \sum_{x \in S} d(x) \qquad Vol(V) = 2m$$

Conductance of S

$$\Phi(S) = \frac{|\partial(S)|}{\min(vol(S), vol(\bar{S}))}$$

Given $p \equiv 1 \times n$ vectors

prob/deg ordering $v_1 \cdots v_{N_p}$ st $N_p = |Supp(p)|$

$$\frac{p(v_i)}{d(v_i)} \geq \frac{p(v_{i+1})}{d(v_{i+1})}$$

---

sweep-set $\equiv S_j^p = \{v_1, \ldots, v_j\} \quad j \leq N_p$

Smallest sweep set cond

$$\Phi(p) = \min_{j \in [1, \ldots, N_p]} \Phi(S_j^p)$$

Observation:

$$pr_\alpha(s)W = pr_\alpha(sW)$$

---

pf $pr_\alpha(s) = \alpha s + (1-\alpha) pr_\alpha(s) W$

$$pr_\alpha(s) W = \alpha s W + (1-\alpha) pr_\alpha(s) W^2$$

$$(pr_\alpha(s) W) = \alpha (sW) + (1-\alpha)(pr_\alpha(s) W) W$$

Thus $pr_\alpha(s) = \alpha s + (1-\alpha) pr_\alpha(sW)$

Def ε-approx PageRank for $pr_\alpha(s)$ is

$pr_\alpha(s-r)$ $r \geq 0$ $r(v) \leq \varepsilon d(v)$

ie $pr_\alpha(s')$ st $s-s' \geq 0$ & $s-s' \leq \varepsilon d$

<u>push (u)</u>

input: p & r

1) $p' := p$ & $r' := r$

2) $p'(u) = p(u) + \alpha r(u)$

3) $r'(u) = (1-\alpha) r(u) / 2$

4) $\forall (u,v) \in E$

$r'(v) = r(v) + (1-\alpha) r(u) / 2d(u)$

---

ie) Push (u)

$p' \Leftarrow p + \alpha r(u) \chi_u$

$r' \Leftarrow r - r(u) \chi_u + (1-\alpha) r(u) \chi_u W.$

<u>Lemma</u>

$$p = pr_\alpha(s-r) \implies p' = pr_\alpha(s-r')$$

Pf

$$pr_\alpha(r) = pr_\alpha(r - r(u)\chi_u) + pr_\alpha(r(u)\chi_u)$$

$$= pr_\alpha(r - r(u)\chi_u) + \alpha r(u)\chi_u + pr_\alpha((1-\alpha)r(u)\chi_u W)$$

$$= pr_\alpha(r - r(u)\chi_u + (1-\alpha)r(u)\chi_u W) + \alpha r(u)\chi_u$$

$$= pr_\alpha(r') + p' - p$$

$$\cancel{pr_\alpha(r)} = pr_\alpha(r') + p' - pr_\alpha(s-r)$$

$$- pr_\alpha(s) + \cancel{pr_\alpha(r)}$$

$$p' = pr_\alpha(s) - pr_\alpha(r') = pr_\alpha(s - r') \quad \square$$

Alg ApproxPR(s, α, ε)

1) Set $p = \vec{0}$ & $r = s$
2) While $r(u) \geq \varepsilon d(u)$ some $u \in V$
   a) push(u)
3) Return p & r

Thm $p' = \text{ApproxPR}(s, \alpha, \varepsilon)$ for $\|s\|_1 \le 1$ & $0 < \varepsilon \le 1$

1) p' is an ε-approx to $pr_\alpha(s)$
2) running time is $O\left(\frac{1}{\varepsilon\alpha}\right)$
3) $\text{vol}(\text{Supp}(p)) \le \frac{2}{(1-\alpha)\varepsilon}$

Pf 1) Def

2) Let T = # pushes

Let d_i = degree of ith push say u

before $r_i(u) \ge \varepsilon d(u)$ $\quad r = r_i$

$r' = r_{i+1}$

$\|r\|_1 - \|r'\|_1$

$$r(u) - r'(u) = r(u)\left(1 - \frac{1-\alpha}{2}\right) = \frac{(1+\alpha)}{2} r(u)$$

$$r(v) - r'(v) = -\frac{(1-\alpha)}{2d(u)} r(u)$$

$$\sum_v r(v) - r'(v) = -\frac{(1-\alpha)}{2} r(u)$$

$$\|r\|_1 - \|r'\|_1 = \frac{(1+\alpha)}{2} r(u) - \frac{(1-\alpha)}{2} r(u)$$

$$= \alpha r(u)$$

$$\geq \varepsilon \alpha d_i$$

Since $\|r_0\|_1 \leq 1$ $\quad \alpha \varepsilon \sum_{i=1}^{T} d_i \leq 1$

$$\sum_{i=1}^{T} d_i \leq \frac{1}{\varepsilon \alpha} \quad \text{work time } i \simeq d_i$$

$O\left(\frac{1}{\varepsilon \alpha}\right)$ work

---

3) Support?

Def $u \in$ Push if $\exists$ push(u) done.

$$|\mathrm{Supp}(p)| \leq \sum_{u \in \mathrm{Push}} d(u)$$

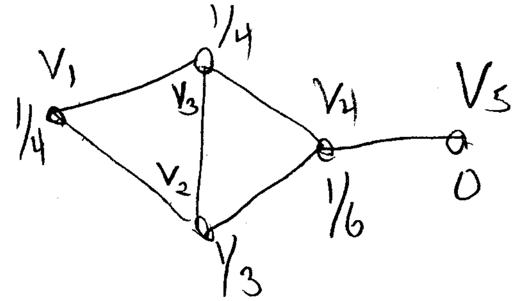
$u \in$ Push 1) $r(u) \geq \left(\frac{1-\alpha}{2}\right) \varepsilon d(u)$

2) $\displaystyle\sum_{u \in \mathrm{Push}} \left(\frac{1-\alpha}{2}\right) \varepsilon d(u) \leq 1$

$$|\mathrm{Supp}(p)| \leq \sum_{u \in P_{\mathrm{vah}}} d(u) \leq \frac{2}{(1-\alpha)\varepsilon}$$

Lovasz-Simonovits Curve

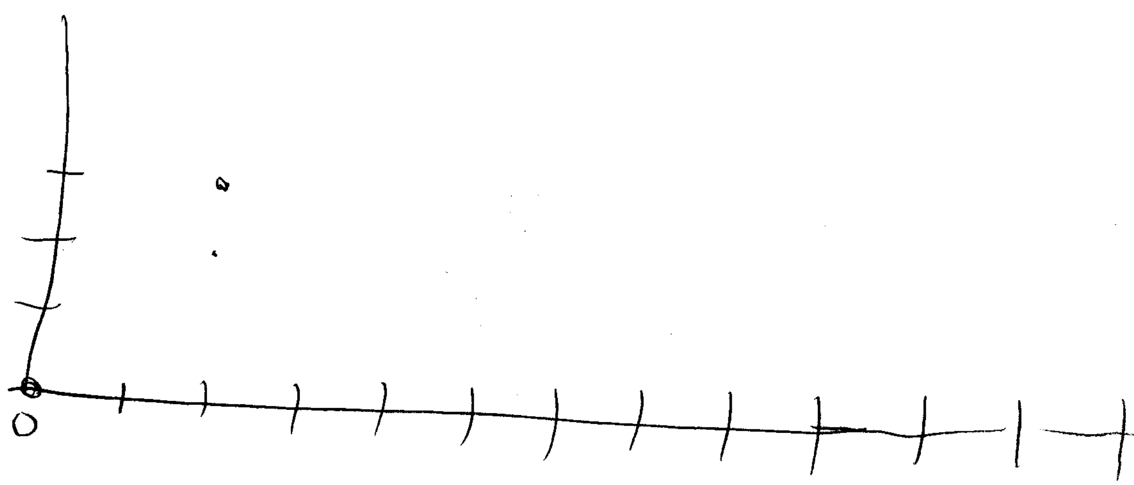
G p

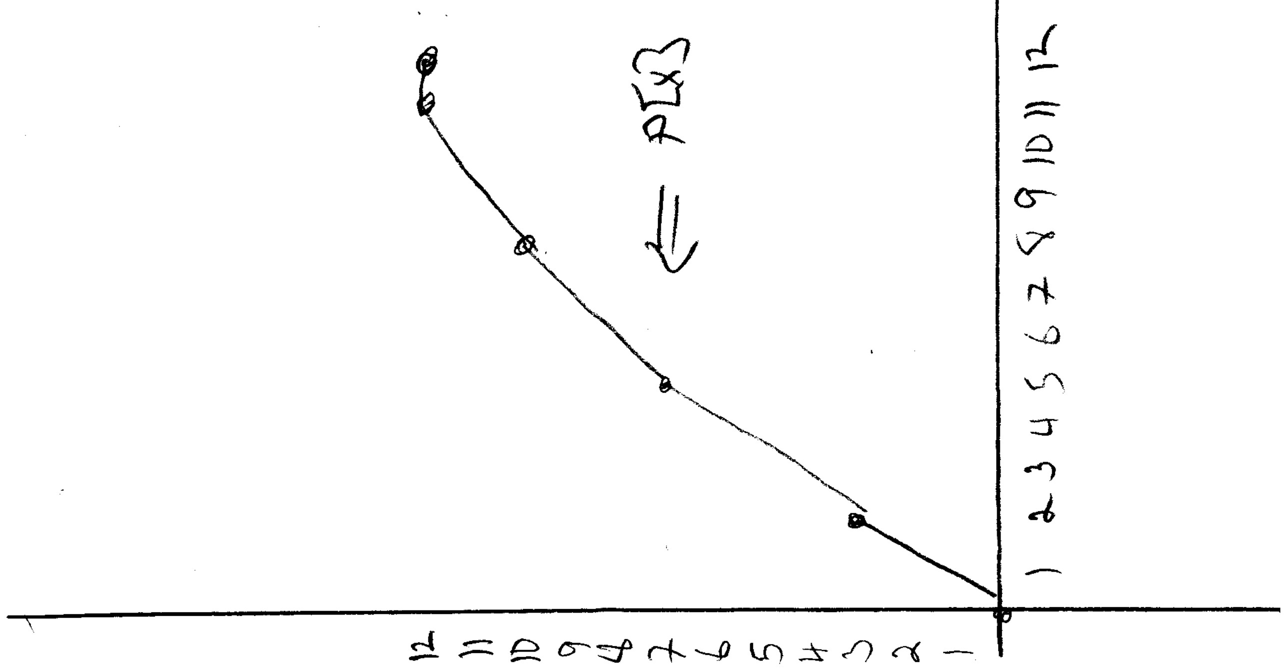


prob/deg	$\frac{1}{8}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	0
j	1	2	3	4	5
$vol(S_j^p)$	2	5	8	11	12
$p(S_j^p)$	$\frac{3}{12}$	$\frac{7}{12}$	$\frac{10}{12}$	$\frac{12}{12}$	$\frac{12}{12}$

$p[2] = 1/4 \quad p[5] = \frac{7}{12}$

$p[x]$




1 2 3 4 5 6 7 8 9 10 11 12
PEx3
1 2 3 4 5 6 7 8 9 10 11 12