Random Walks, Mixing Time on Iterative Solvers

Random Walks

$G = (V, E, w) \quad A_{ij} = w_{ij}$

$D_{ii} = \sum_j w_{ij}$

Let $p_i^{(t)} \equiv$ prob at vertex V_i at time t.

Claim $A^T D^{-1} p^{(t)} = p^{(t+1)}$

note prob go from V_i to V_j $= \dfrac{w_{ij}}{d_i}$ $\quad d_i = \sum_j w_{ij}$

$$\left(\underset{i}{\overset{A^T}{\begin{bmatrix} w_{i1} \\ \vdots \\ w_{in} \end{bmatrix}}} \begin{bmatrix} & & \\ & d_i^{-1} & \\ & & \end{bmatrix} \right)_i = \begin{pmatrix} w_{i1}/d_i \\ \vdots \\ w_{in}/d_i \end{pmatrix}$$

Note Col sums in $A^T D^{-1}$ are 1.

$A^T D^{-1}$ called transition matrix

In our case $A = A^T \wedge \sum_i p_i^{(0)} = 1 \quad p_i^{(0)} \geq 0$

Two natural questions

1) $\exists$ dist $\bar{P}$ s.t. $AD^{-1}\bar{P} = \bar{P}$ (stationary dist)

2) $\forall P_0$ $\lim\limits_{k\to\infty} (AD^{-1}) P_0 = \bar{P}$

1) Yes let $d = \sum d_i$ $\quad \pi = \begin{pmatrix} d_1/d \\ \vdots \\ d_n/d \end{pmatrix} = 1/d \, D \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$

$$AD^{-1}\pi = AD^{-1}(1/d) D \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = 1/d \, A \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = 1/d \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = \pi$$

$\bar{d} = \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix}$ is an eigenvector with value 1

2) In general no. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ for $G \equiv \circ\!\!-\!\!\bullet$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Note no iff $\exists x$ st $AD^{-1}x = -x$

$\lambda = -1$

<u>Thm</u> If G is not bipartite & connected then

$$\forall p^{(0)} \; |p^{(0)}|_1 = 1 \;\&\; p^{(0)} \geq 0 \quad \lim_{k \to \infty} (AD^{-1}) p^{(0)} = \pi$$

<u>Question</u> How fast does G "mix"?

<u>Prob</u> A sym but AD^{-1} is <u>not</u>!

We do a change of variables.

$$AD^{-1} \longrightarrow \tilde{A} = D^{-1/2} A D^{-1/2} \qquad\qquad y = D^{-1/2} x$$

$$p^{(k)} \longrightarrow \tilde{p}^{(k)} = D^{-1/2} p^{(k)}$$

$$\pi \longrightarrow \tilde{\pi} = D^{-1/2} \pi$$

<u>Claim</u> $AD^{-1} x = \lambda x$ iff $\tilde{A} y = \lambda \tilde{y}$

$(\Rightarrow)$ $y = D^{-1/2} x$

$$\tilde{A} y = D^{-1/2} A D^{-1/2} D^{-1/2} x = D^{-1/2} A D^{-1} x = \lambda D^{-1/2} x = \lambda y$$

Spectral Thm

If A is real sym matrix then

1) Eigenvalues of A are real

ie $Ax = \lambda x \Rightarrow \lambda$ is real

2) If $Ax = \lambda x$ & $Ay = \mu y$ & $\lambda \neq \mu$ then

$x^T y = 0$ ie $x \perp y$

3) $\exists$ an orthonormal bases $y_1 \cdots y_n$ st

$$A = \begin{pmatrix} | & & | \\ y_1 & \cdots & y_n \\ | & & | \end{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} - & y_1^T & - \\ & \vdots & \\ - & y_n^T & - \end{pmatrix}$$

4) $A = \sum_i \lambda_i (y_i y_i^T)$

Thus $\tilde{A}\tilde{\pi} = \tilde{\pi}$

Def Error $\quad \varepsilon_k = \tilde{\pi} - \tilde{A}^k \tilde{P}_0 = \tilde{\pi} - \tilde{p}^{(k)}$

Question How fast does ε_k go to 0 with k?

Claim $\varepsilon_0 \perp \tilde{\pi}$ ie $\varepsilon_0^T \tilde{\pi} = 0$

$$\tilde{\pi} = D^{-1/2}\pi = \begin{pmatrix} 1/\sqrt{d_1} & & \\ & \ddots & \\ & & 1/\sqrt{d_n} \end{pmatrix} \begin{pmatrix} d_1/d \\ \vdots \\ d_n/d \end{pmatrix} = \begin{pmatrix} \sqrt{d_1}/d \\ \vdots \\ \sqrt{d_n}/d \end{pmatrix}$$

$$\tilde{\pi}^T \tilde{\pi} = \sum \left(\sqrt{d_i}/d\right)^2 = \sum d_i/d^2 = 1/d$$

$$\tilde{\pi}^T p^{(0)} = \left(\sqrt{d_1}/d \cdots \sqrt{d_n}/d\right) \begin{pmatrix} p_1/\sqrt{d_1} \\ \vdots \\ p_n/\sqrt{d_n} \end{pmatrix} = \sum p_i/d = 1/d \sum p_i = 1/d$$

$$\tilde{\pi}^T \varepsilon_0 = \tilde{\pi}^T\left(\tilde{\pi} - \tilde{p}^{(0)}\right) = \tilde{\pi}^T\tilde{\pi} - \tilde{\pi}^T \tilde{p}^{(0)} = d - d = 0$$

Suppose eigenvalues of $\tilde{A}$

$$-1 < \lambda_1 \le \cdots \le \lambda_n = 1$$

Eigenvectors $Y_1 \cdots Y_n = \tilde{\pi} \cdot d$

We know $\mathcal{E}_0 = \alpha_1 Y_1 + \cdots + \alpha_{n-1} Y_{n-1}$

$$\mathcal{E}_1 = \tilde{A}\,\mathcal{E}_0 = \lambda_1 \alpha_1 Y_1 + \cdots \quad \lambda_{n-1} \alpha_{n-1} Y_{n-1}$$

$$\mathcal{E}_k = \lambda_1^k \alpha_1 Y_1 + \quad + \lambda_{n-1}^k \alpha_{n-1} Y_{n-1}$$

Thus mixing rate is determined by

$$\lambda = \max\{|\lambda_1|, |\lambda_{n-1}|\}$$

Pick k s.t. $\lambda^k \le 1/2$

Thus every k rounds the error halves.

The Basic Iterative Method

Goal: Solve $Au=b$ (*) A arbitrary

$u^{(0)} \equiv$ initial guess

Richardson's Method

$$u^{(m+1)} = (I-A)u^{(m)} + b = u^{(m)} + (b - Au^{(m)})$$

Good: 1) Each step is fast & parallel $O(w(A))$ work

2) A fixed point is a solution to (*)

pf $x = (I-A)x + b \Rightarrow 0 = -Ax + b \Rightarrow Ax = b$

Bad: 1) May not converge

2) If it does it may do so very slowly

The error for Richardson

$G = (I - A)$ then

$$x^{(m+1)} = G x^{(m)} + b$$

Suppose $A\bar{u} = b$

error $\varepsilon^{(m)} = u^{(m)} - \bar{u}$

<u>Claim</u> $\varepsilon^{(m)} = G^m \varepsilon^{(0)}$

<u>to show</u> $\varepsilon^{(m+1)} = G \varepsilon^{(m)}$

<u>pf</u>
$$\begin{aligned} G\varepsilon^{(m)} &= G u^{(m)} - (I - A)\bar{u} \\ &= G u^{(m)} - \bar{u} + A\bar{u} \\ &= G u^{(m)} - \bar{u} + b \\ &= \underbrace{G u^{(m)} + b} - \bar{u} \\ &= u^{(m+1)} - \bar{u} = \varepsilon^{(m+1)} \end{aligned}$$

We need that $\lim_{m\to\infty} G^m \varepsilon^0 = 0 \quad \forall \varepsilon^0$

ie $\lim_{m\to\infty} G^m = 0$

By Spectral Thm

$$G = \sum_{i=1}^{n} \lambda_i (v_i v_i^T) \quad \& \quad G^k = \sum \lambda_i^k v_i v_i^T$$

Thus $\lim_{m\to\infty} G^m = 0$ iff $-1 < \lambda_1 \cdots \lambda_n < 1$

<u>Def</u> $\lambda(G) \equiv$ eigen values of G

$\lambda_{Max}(G) \equiv$ max eigenvalues

$\lambda_{Min}(G) \equiv$ min eigen

Extrapolated method

new recurrence

$$u^{(m+1)} = u^{(m)} + \gamma\,(\underbrace{b - A u^{(m)}}_{\text{residual error}})$$

$$= (I - \gamma A)\, u^{(m)} + \gamma b$$

$$G_\gamma = (I - \gamma A)$$

Goal: Find $\min_{\gamma} \max \{ |\lambda_{max} G_\gamma|, |\lambda_{min} G_\gamma| \}$

Assume A spd

$$0 < m = \lambda_{min}(A) < M = \lambda_{max}(A)$$

pick $\gamma = \dfrac{2}{M+m}$

$$\lambda_{max}(G_\gamma) = \lambda_{max}(I - \gamma A) = 1 - \left(\frac{2}{M+m}\right)m$$

$$= \frac{M+m-2m}{M+m} = \frac{M-m}{M+m} < 1$$

$$\lambda_{min}(G_\gamma) = 1 - \left(\frac{2}{M+m}\right)M = \frac{m-M}{M+m} = -\left(\frac{M-m}{M+m}\right) > -1$$

<u>Def</u> Condition number of $A \equiv \frac{M}{m} \equiv K$

$$\lambda_{Max}(G_\gamma) = \frac{M-m}{M+m} = \frac{M/m - 1}{M/m + 1} = \frac{K-1}{K+1} = 1 - 1/K$$

<u>Thus</u> $|\lambda(G_\gamma)| \leq 1 - 1/K$

$$|\lambda(G_\gamma^K)| \leq (1 - 1/K)^K \doteq \frac{1}{e}$$

Error halves every K iterations

What does this mean?

Graph $\equiv L_n$ line graph L_{L_n}

Claim $\lambda(L) \equiv 0 = \lambda_1 < \lambda_2 \cdots \leq \lambda_n \doteq 2$

$\lambda_2 \doteq 1/n^2$

ignore λ_1 $\quad k(L) \approx n^2$

Thus Richardson's needs n^2 iteration per bit.