15-859N
9/13/13

Random Walks, Mixing Times

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 going from V_i to $V_j = \dfrac{w_{ij}}{d_i}$ $\quad d_i = \sum_j w_{ij}$

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

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) $\forall P_0 \quad \lim_{k\to\infty} (AD^{-1})^k P^{(0)} = \bar{P}$

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

<u>Pf</u>

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

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

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

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

<u>Note</u> no iff $\exists x$ st $AD^{-1}x = -x$

$\lambda = -1$

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

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

<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 $\tilde{A} y = \lambda y$

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

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

$(\Leftarrow)$

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 (eigenvectors)

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

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

See page 10

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

<u>Def</u> Error $\tilde{\varepsilon}^{(k)} = \tilde{\pi} - \tilde{A}^{k}\tilde{p}^{(0)} = \tilde{\pi} - \tilde{p}^{(k)}$

<u>Question</u> How fast does ε_k go to 0 with k?

<u>Claim</u> $\tilde{\varepsilon}_0 \perp \tilde{\pi}$ ie $\tilde{\varepsilon}_0^T \tilde{\pi} = 0$

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

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

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

$$\tilde{\pi}^T\tilde{\varepsilon}^{(0)} = \tilde{\pi}^T\left(\tilde{\pi} - \tilde{p}^{(0)}\right) = \tilde{\pi}^T\tilde{\pi} - \tilde{\pi}^T\tilde{p}^{(0)} = \frac{1}{d} - \frac{1}{d} = 0$$

Perron-Frobenius Thm

Suppose $A^{n \times n} \geq 0$ Graph(A) is strongly connected

Def $z \in \mathbb{C}$ $|z| = \sqrt{z^* z} = \sqrt{a^2 + b^2}$ $z = a + ib$

Spectral radius $\rho(A) = \max\limits_{\lambda \in \lambda(A)} |\lambda|$

Thm 1) $\rho(A)$ is a simple eigenvalue of A.

If x is an eigenvector for ρ then $\text{sign}(x_i) = \text{sign}(X_j)$ $\forall i, j$

2) $\Theta \in \lambda(A)$ and $|\Theta| = \rho(A)$ then $\Theta / \rho(A)$ is an mth root of unit and all cycles in X have length a multiple of m.

3) Only non-neg eigenvector is X.

proof (to come)

Suppose eigenvalues of $\tilde{A}$ are:

$$-1 < \lambda_1 \leq \cdots \leq \lambda_n = 1 \quad \text{with}$$

Eigenvectors $Y_1, \ldots, Y_n = \tilde{\Pi} \cdot d$

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

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

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

Thus mixing rate is determined by

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

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

Thus every k rounds the error halves.

Pf $\|\tilde{\varepsilon}^{(k)}\|_2 = \sqrt{\sum \lambda_i^{2k} \alpha_i^2 \lambda_i^2}$

$$\le \sqrt{\sum \lambda^{2k} \alpha_i^2 \lambda_i^2}$$

$$= \lambda^k \sqrt{\sum \alpha_i^2 \lambda_i^2}$$

$$= \lambda^k \|\tilde{\varepsilon}^{(0)}\|_2$$

$\lambda^k \le 1/2$ done.

PROOF OF SPECTRAL THEOREM

Theorem 1 (Spectral Theorem). *Let $A \in \mathbb{R}^{n\times n}$ be a real symmetric matrix. Then*

(1) *All eigenvalues of A are real.*
(2) *There exists an orthogonal matrix Q and a diagonal matrix Λ such that $A = Q\Lambda Q^T$.*

Proof. We have proved (1) in the class. Only need to prove (2). We make an induction on n.

When $n = 1$, the claim is obvious. Now assume that the claim is valid for $n = m$, that is, for any $m \times m$-real symmetric matrix A, there exists an orthogonal matrix Q and diagonal matrix Λ such that $A = Q\Lambda Q^T$. Let us consider $(m+1) \times (m+1)$-real symmetric matrix A. By (1), A has a real eigenvalue λ with eigenvector α. We see that all entries of α must be real numbers. By Gram-Schmidt process, we may assume that there exists an *orthonormal* basis $q_1, \dots, q_n$ with $q_1 = \alpha$. Let $P := (q_1 q_2 \cdots q_n)$ and $C := P^T AP = (c_{ij})_{(m+1)\times(m+1)}$. We claim that $c_{11} = \lambda$ and $c_{i1} = 0$ for $i \neq 1$. In fact, note that P is an orthogonal matrix, we have $AP = PC$, that is, $A(q_1 q_2 \cdots q_n) = (q_1 q_2 \cdots q_n)C$. Therefore, we have $Aq_1 = \sum_{i=1}^{m+1} c_{i1} q_i$. But $q_1 = \alpha$ is an eigenvector,so $\lambda q_1 = \sum_{i=1}^{m+1} c_{i1} q_i$. Since $q_1, \cdots, q_n$ are linearly independent. So $c_{11} = \lambda$ and $c_{i1} = 0$ for $i \neq 1$. So C has four blocks like $\begin{pmatrix} \lambda & \star \\ 0 & \tilde{A} \end{pmatrix}$. Note that $C = P^T AP$ is symmetric(why ?), thus $\star = 0$. So $C = \begin{pmatrix} \lambda & 0 \\ 0 & \tilde{A} \end{pmatrix}$ and $\tilde{A}$ has to be symmetric matrix with the size $m \times m$. By induction, there exists an orthogonal matrix Q and diagonal matrix Λ such that $\tilde{A} = Q\Lambda Q^T$. Therefore

$$C = \begin{pmatrix} \lambda & 0 \\ 0 & \tilde{A} \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & Q\Lambda Q^T \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \Lambda \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}^T$$

Therefore

$$A = PCP^T = P \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \Lambda \end{pmatrix} (P \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix})^T$$

and we easily check $P \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}$ is an orthogonal matrix and we are done. □