Based on lecture notes by Michael Mahoney, Dan Spielman, Luca Trevisan, etc.

Random walks

- Let μ^t be the probability distribution at time t.
$$\mu^{t+1} = (AD^{-1})\mu^t$$
- $(AD^{-1})(D\mathbf{1}) = A\mathbf{1} = (D\mathbf{1})$.
- Therefore, stationary distribution $\mu_i^* = d_i/\left(\sum_j d_j\right)$.
- AD^{-1} has the same eigenvalues as $D^{-1/2}AD^{-1/2}$.

- Theorem: Let M be a $n \times n$ matrix and let S be a $n \times n$ invertible matrix. Then, M and SMS^{-1} have the same eigenvalues.
- Let v an eigenvector of M with eigenvalue λ. Let $v' \stackrel{\text{def}}{=} Sv$. Then
$$SMS^{-1}v' = SMS^{-1}Sv = S(\lambda v) = \lambda v'$$
- (H.W.) Mixing time bounds for random walks on general graphs.

How small can β be?

- Let G be a d-regular graph, with eigenvalues $d \geq \sigma_2 \geq \cdots \geq \sigma_n$ where $\max_{i \in \{2,\ldots,n\}} |\sigma_i| \leq \beta$.
- $\mathrm{Tr}(A^2) = \sum_i e_i^T A^2 e_i = \sum_i \|Ae_i\|^2 = nd$
- $\mathrm{Tr}\left(A^k\right)$ = number of walks of length k starting and ending at the same vertex.

$$nd = \mathrm{Tr}(A^2) = \sum_i \sigma_i^2 \leq d^2 + (n-1)\beta^2$$

- Therefore, $\beta \geq \sqrt{\frac{n-d}{n-1}} \cdot \sqrt{d} = \left(1 - o(1)\right)\sqrt{d}$
- It can be shown that $\beta \geq 2\sqrt{d-1} - o(1)$

Ramanujan Graphs

- Ramanujan graph: A d-regular graph is called a Ramanujan graph if $\beta \leq 2\sqrt{d-1}$
- For a random n-vertex d-regular graph, $\beta \leq 2\sqrt{d-1} + o(1)$ w.h.p. [Friedman - 03].
- Explicit constructions known for some values of d [Lubotzky, Philips, Sarnak - 88], etc.

Expander mixing lemma

- Theorem: Let G be a d-regular graph on n vertices. Then for any $S, T \subset V$
$$\left| |E(S,T)| - \frac{d}{n}|S||T| \right| \leq \beta\sqrt{|S||T|}$$

- For $S \subseteq V$ and let $\mathbf{1}_S$ be its indicator vector. Let $\mathbf{1}_S = c_S\mathbf{1} + p_S$ where $\langle p_S, \mathbf{1}\rangle = 0$. Then
$$\langle \mathbf{1}_S, \mathbf{1}\rangle = \langle c_S\mathbf{1}, \mathbf{1}\rangle + \langle p_S, \mathbf{1}\rangle = c_S n + 0$$
- Fix $S, T \subset V$, $\mathbf{1}_S = (|S|/n)\mathbf{1} + p_S$ and $\mathbf{1}_T = (|T|/n)\mathbf{1} + p_T$.
$$|E(S,T)| = \mathbf{1}_S^T A \mathbf{1}_T = \left(\frac{|S|}{n}\mathbf{1} + p_S\right)^T A \left(\frac{|T|}{n}\mathbf{1} + p_T\right) = \left(\frac{|S|}{n}\mathbf{1} + p_S\right)^T \left(d\frac{|T|}{n}\mathbf{1} + Ap_T\right)$$
$$= d\frac{|S|}{n}\frac{|T|}{n}\mathbf{1}^T\mathbf{1} + \frac{|S|}{n}\mathbf{1}^T Ap_T + d\frac{|T|}{n}p_S^T\mathbf{1} + p_S^T Ap_T = d\frac{|S||T|}{n} + p_S^T Ap_T$$

- Therefore,

$$\left| |E(S,T)| - \frac{d}{n}|S||T| \right| = \left| p_S^T A p_T \right| \leq \|p_S\| \|Ap_T\|$$

- $\|Ap_T\| = \sqrt{p_T^T A^T A p_T} \leq \sqrt{\beta^2 p_T^T p_T} = \beta \|p_T\|$

$$\left| |E(S,T)| - \frac{d}{n}|S||T| \right| \leq \|p_S\| \|Ap_T\| \leq \beta \|p_S\| \|p_T\| \leq \beta \|\mathbf{1}_S\| \|\mathbf{1}_T\| = \beta \sqrt{|S||T|}$$

Laplacian Matrix

- $L = D - A$ and $L\mathbf{1} = 0$.
- $\mathcal{L} = D^{-1/2} L D^{-1/2}$ and $\mathcal{L}\left(D^{1/2}\mathbf{1}\right) = D^{-1/2} L D^{-1/2}\left(D^{1/2}\mathbf{1}\right) = D^{-1/2} L\mathbf{1} = 0$

$$x^T L x = x^T (D - A) x = \sum_i d_i x_i^2 - 2\sum_{ij\in \mathrm{E}} A_{ij} x_i x_j = \sum_i \left(\sum_j A_{ij}\right) x_i^2 - 2\sum_{ij\in \mathrm{E}} A_{ij} x_i x_j$$

$$= \sum_{ij\in E} A_{ij}\left(x_i - x_j\right)^2$$

$$\lambda_2 = \min_{x\perp D^{1/2}\mathbf{1}} \frac{x^T \mathcal{L} x}{x^T x} = \min_{x\perp D^{1/2}\mathbf{1}} \frac{\left(D^{-1/2}x\right)^T L\left(D^{-1/2}x\right)}{x^T x} = \min_{y\perp D\mathbf{1}} \frac{y^T L y}{y^T D y} = \min_{y\perp D\mathbf{1}} \frac{\sum_{ij\in E} A_{ij}\left(y_i - y_j\right)^2}{\sum_i d_i y_i^2}$$

- If $y = D^{-1/2}x$, then $\left\langle x, D^{1/2}\mathbf{1}\right\rangle = \langle y, D\mathbf{1}\rangle$.

Laplacian Matrix

- Let $x \in \{0,1\}^n$ and let $S_x = \{i: x_i \neq 0\}$.
- For a pair of vertices i and j, $\left(x_i - x_j\right)^2$ indicates whether i and j are separated by S_x.

$$\frac{\sum_{ij \in E} A_{ij}\left(x_i - x_j\right)^2}{\sum_i d_i x_i^2} = \frac{\sum_{i \in S_x, j \in V \setminus S_x} A_{ij}}{\sum_{i \in S_x} d_i}$$

- Therefore,

$$\phi_G = \min_{S: \mathrm{vol}(S) \leq \mathrm{vol}(V)/2} \phi(S) = \min_{x \in \{0,1\}^n:\, \langle x, D\mathbf{1} \rangle \leq \sum_i d_i/2} \frac{\sum_{ij \in E} A_{ij}\left(x_i - x_j\right)^2}{\sum_i d_i x_i^2}$$

- $\langle x, D\mathbf{1} \rangle = \sum_i d_i x_i = \mathrm{vol}(S_x)$

$\lambda_2/2 \leq \phi_G$

$$\lambda_2 = \min_{y \perp D\mathbf{1}} \frac{\sum_{ij \in E} A_{ij}(y_i - y_j)^2}{\sum_i d_i y_i^2}$$

- Fix any $S \subset V$ such that $\mathrm{vol}(S) \leq \mathrm{vol}(V)/2$, and let x be the indicator vector of S, i.e. $x_i = \begin{cases} 1 \ if \ i \in S \\ 0 \ if \ i \notin S \end{cases}$
- Let $y = x + c\mathbf{1}$ such that $\langle y, D\mathbf{1} \rangle = 0$.

$$0 = \langle y, D\mathbf{1} \rangle = \langle x, D\mathbf{1} \rangle + c\langle \mathbf{1}, D\mathbf{1} \rangle = \sum_i d_i x_i + c \sum_i d_i$$

- Therefore, $c = -\mathrm{vol}(S) \,/\, \mathrm{vol}(V)$.
- Note that for any i, j, we have $y_i - y_j = x_i - x_j$. Therefore,

$$\sum_{ij \in E} A_{ij}(y_i - y_j)^2 = \sum_{ij \in E} A_{ij}(x_i - x_j)^2 = \sum_{i \in S, j \in V \setminus S} A_{ij}$$

$$\sum_i d_i y_i^2 = \sum_i d_i (x_i + c)^2 = \sum_i d_i x_i^2 + c^2 \sum_i d_i + 2c \sum_i d_i x_i$$
$$= \mathrm{vol}(S) + \left(\frac{\mathrm{vol}(S)}{\mathrm{vol}(V)}\right)^2 \mathrm{vol}(V) \; - 2\frac{\mathrm{vol}(S)}{\mathrm{vol}(V)}\mathrm{vol}(S) = \mathrm{vol}(S)\left(1 - \frac{\mathrm{vol}(S)}{\mathrm{vol}(V)}\right)$$
$$\geq \frac{1}{2}\mathrm{vol}(S)$$

- Therefore,

$$\lambda_2 \leq \frac{\sum_{ij\in E} A_{ij}\left(y_i - y_j\right)^2}{\sum_i d_i y_i^2} \leq 2\frac{\sum_{i\in S, j\in V\setminus S} A_{ij}}{\mathrm{vol}(S)} = 2\phi(S)$$

- Lemma 1: There exists a polynomial time algorithm that takes a graph $G = (V, E)$ and an $x \in \mathbb{R}^n_{\geq 0}$, and computes an $S \subseteq \mathrm{supp}(x)$ such that

$$\frac{\sum_{i\in S, j\in V\setminus S} A_{ij}}{\sum_{i\in S} d_i} \leq \frac{\sum_{ij} A_{ij}\left|x_i - x_j\right|}{\sum_i d_i x_i}$$

- Without loss of generality assume that $x_1 \geq x_2 \geq \cdots \geq x_n \geq 0$.
- For $i < j$, writing $x_i - x_j = \sum_{l=i}^{j-1}(x_l - x_{l+1})$ and $x_i = \sum_{l=i}^{n}(x_l - x_{l+1})$ where $x_{n+1} \stackrel{\text{def}}{=} 0$.

$$\frac{\sum_{ij} A_{ij}\left|x_i - x_j\right|}{\sum_i d_i x_i} = \frac{\sum_{i<j} A_{ij}\sum_{l=i}^{j-1}(x_l - x_{l+1})}{\sum_i d_i \sum_{l=i}^{n}(x_l - x_{l+1})} = \frac{\sum_{l=1}^{n}\left((x_l - x_{l+1})\sum_{i\in[l], j\in V\setminus[l]} A_{ij}\right)}{\sum_{l=1}^{n}\left((x_l - x_{l+1})\sum_{i\in[l]} d_i\right)}$$

- $(x_l - x_{l+1})$ will appear for ij when $i \leq l$ and $j \geq l + 1$.

- Fact: For $a_1, \ldots, a_n, b_1, \ldots, b_n, c_1, \ldots, c_n > 0$, we have

$$\frac{c_1 a_1 + \cdots + c_n a_n}{c_1 b_1 + \cdots c_n b_n} \geq \min_i \frac{a_i}{b_i}$$

- Let $\alpha = \min_i a_i / b_i$. Then $a_i \geq \alpha b_i \ \forall i$. Therefore,

$$\frac{c_1 a_1 + \cdots + c_n a_n}{c_1 b_1 + \cdots c_n b_n} \geq \frac{c_1 \alpha b_1 + \cdots + c_n \alpha b_n}{c_1 b_1 + \cdots c_n b_n} = \alpha$$

$$\frac{\sum_{ij} A_{ij} |x_i - x_j|}{\sum_i d_i x_i} = \frac{\sum_{l=1}^{n} \left((x_l - x_{l+1}) \sum_{i \in [l], j \in V \setminus [l]} A_{ij} \right)}{\sum_{l=1}^{n} \left((x_l - x_{l+1}) \sum_{i \in [l]} d_i \right)} \geq \min_{l : (x_l - x_{l+1}) > 0} \frac{\sum_{i \in [l], j \in V \setminus [l]} A_{ij}}{\sum_{i \in [l]} d_i}$$

- Therefore, $S = [l^*]$ for optimal l^* above suffices. Since $(x_{l^*} - x_{l^*+1}) > 0$, $S \subseteq \operatorname{supp}(x)$.

- Lemma 2: There exists a polynomial time algorithm that takes a graph $G = (V, E)$ and a $y \in \mathbb{R}^n_{\geq 0}$, and computes an $S \subseteq \mathrm{supp}(y)$ such that

$$\frac{\sum_{i \in S, j \in V \setminus S} A_{ij}}{\sum_{i \in S} d_i} \leq \sqrt{2\frac{y^T L y}{y^T D y}}$$

- Idea: Use Lemma 1 with x, where $x_i \stackrel{\text{def}}{=} y_i^2$. Note that $\mathrm{supp}(x) = \mathrm{supp}(y)$

$$\sum_{ij} A_{ij}|x_i - x_j| = \sum_{ij} A_{ij}|y_i - y_j|(y_i + y_j) \leq \sqrt{\sum_{ij} A_{ij}(y_i - y_j)^2} \sqrt{\sum_{ij} A_{ij}(y_i + y_j)^2}$$

$$= \sqrt{y^T L y}\sqrt{\sum_{ij} A_{ij}(y_i^2 + y_j^2 + 2y_i y_j)} \leq \sqrt{y^T L y}\sqrt{2\sum_{ij} A_{ij}(y_i^2 + y_j^2)} = \sqrt{y^T L y}\sqrt{2\sum_i d_i y_i^2}$$

- Therefore,

$$\frac{\sum_{ij} A_{ij}|x_i - x_j|}{\sum_i d_i x_i} \leq \frac{\sqrt{y^T L y}\sqrt{2\sum_i d_i y_i^2}}{\sum_i d_i y_i^2} = \sqrt{2\frac{y^T L y}{y^T D y}}$$

- Lemma 3: There exists a polynomial time algorithm that takes a graph $G = (V, E)$ and a $z \in \mathbb{R}^n$ such that $\langle z, D\mathbf{1} \rangle = 0$, and computes an S such that $\mathrm{vol}(S) \leq \mathrm{vol}(V)/2$ and $\phi(S) \leq \sqrt{2\frac{z^T L z}{z^T D z}}$.

- Idea: (1) Shift every entry of z such that volume of the support of the positive part of z and the negative part of z is at most half of the total volume. (2) Use Lemma 2 on only the positive part or the negative part, whichever is "better".

- For $a \in \mathbb{R}$, let $a^+ \stackrel{\text{def}}{=} \max\{a, 0\}$ and $a^- \stackrel{\text{def}}{=} \max\{-a, 0\}$. Note that a $= a^+ - a^-$.
- Let $u = z + c\mathbf{1}$ for an appropriate constant c such that $\mathrm{vol}(\mathrm{supp}(u^+)), \mathrm{vol}(\mathrm{supp}(u^-)) \leq \mathrm{vol}(V)/2$
- $\sum_{ij} A_{ij}(z_i - z_j)^2 = \sum_{ij} A_{ij}(u_i - u_j)^2$
- $\sum_i d_i u_i^2 = \sum_i d_i (z_i + c)^2 = \sum_i d_i z_i^2 + c^2 \sum_i d_i + 2\sum_i d_i z_i \geq \sum_i d_i z_i^2$.

$$\frac{\sum_{ij} A_{ij}(u_i - u_j)^2}{\sum_i d_i u_i^2} \leq \frac{\sum_{ij} A_{ij}(z_i - z_j)^2}{\sum_i d_i z_i^2}$$

- Claim: $(a-b)^2 \geq (a^+ - b^+)^2 + (a^- - b^-)^2$.

If $a, b \geq 0$, then $a^+ = a, b^+ = b, a^- = b^- = 0$.

If $a \geq 0, b < 0$, then $a^+ = a, b^+ = 0, a^- = 0, b^- = -b$. Therefore,

$$(a-b)^2 = a^2 + b^2 - 2ab \geq a^2 + b^2 = (a^+ - b^+)^2 + (a^- - b^-)^2$$

- Therefore, $\sum_{ij\in E} A_{ij}\left(u_i - u_j\right)^2 \geq \sum_{ij\in E} A_{ij}\left(u_i^+ - u_j^+\right)^2 + \sum_{ij\in E} A_{ij}\left(u_i^- - u_j^-\right)^2$
- $\sum_i d_i u_i^2 = \sum_i d_i (u_i^+)^2 + \sum_i d_i (u_i^-)^2$

$$\frac{\sum_{ij\in E} A_{ij}\left(u_i - u_j\right)^2}{\sum_i d_i u_i^2} \geq \frac{\sum_{ij\in E} A_{ij}\left(u_i^+ - u_j^+\right)^2 + \sum_{ij\in E} A_{ij}\left(u_i^- - u_j^-\right)^2}{\sum_i d_i\left(u_i^+\right)^2 + \sum_i d_i\left(u_i^-\right)^2}$$

$$\geq \min\left\{\frac{\sum_{ij\in E} A_{ij}\left(u_i^+ - u_j^+\right)^2}{\sum_i d_i\left(u_i^+\right)^2}, \frac{\sum_{ij\in E} A_{ij}\left(u_i^- - u_j^-\right)^2}{\sum_i d_i\left(u_i^-\right)^2}\right\}$$

- Use Lemma 2 on the minimizer above.

- Without loss of generality, assume that $z_1 \geq \cdots \geq z_n$. Then

$$\min_l \phi([l]) \leq \sqrt{2\frac{z^T L z}{z^T D z}}$$

where $\phi(S) = \frac{\sum_{i \in S, j \in V \setminus S} A_{ij}}{\min\{\mathrm{vol}(S), \mathrm{vol}(V \setminus S)\}}$. (why?)

Cheeger's Inequality

- Use Lemma 3 with $z = D^{-1/2} v_2$. Then $\langle z, D\mathbf{1} \rangle = 0$ and $\frac{z^T L z}{z^T D z} = \lambda_2$.

- Cycle: $\phi_G = 2/n$ and $\lambda_2 = 1 - \cos\frac{2\pi}{n} \approx \Theta\left(\frac{1}{n^2}\right)$. Therefore, $\phi_G = \Theta(\sqrt{\lambda_2})$

Hypercube

- $V = \{-1,1\}^d$ and $\{x, y\} \in E$ if x and y differ in exactly one coordinate.
- $\chi_S \stackrel{\text{def}}{=} \Pi_{i \in S} x_i$, where $S \subseteq [d]$. Therefore, $\chi_S \in \{-1,1\}^{2^d}$
- Theorem: χ_S is an eigenvector of the normalized adjacency matrix with eigenvalue $1 - 2|S|/d$.

- Let $S \neq T$. Then

$$\langle \chi_S, \chi_T \rangle = \sum_{x \in \{-1,1\}^d} (\Pi_{i \in S} x_i \Pi_{i \in T} x_i) = 2^d \operatorname{E}_{x \sim \{-1,1\}^d} \left((\Pi_{i \in S \cap T} x_i^2)(\Pi_{i \in S \Delta T} x_i) \right)$$

$$= 2^d \Pi_{i \in S \Delta T} \operatorname{E}_{x_i \sim \{-1,1\}} x_i = 0$$

- Fix $S \subseteq [d]$. Then for any $x \in \{-1,1\}^d$

$$(A\chi_S)_x = \sum_{y \in N(x)} \chi_S(y) = \sum_{y \in N(x)} \Pi_{j \in S} y_j = \sum_{i \in S} \left(-\Pi_{j \in S} x_j\right) + \sum_{i \notin S} \left(\Pi_{j \in S} x_j\right)$$
$$= (d - 2|S|)\Pi_{j \in S} x_j = (d - 2|S|)\chi_S(x)$$

- Therefore, $A\chi_S = (d - 2|S|)\chi_S$
- Therefore, eigenvalues of $\mathcal{L}$ are $\left\{\frac{2i}{d} : i = 0, \dots, d\right\}$ where $2i/d$ has multiplicity $\binom{d}{i}$.

- Let $S_i = \{x \in V : x_i = 1\}$.

$$\phi_G \leq \phi(S_i) = \frac{1}{d} = \frac{\lambda_2}{2}$$