Random Walks on Graphs

Following Spielman, Chapter 10. For the randomness extractors part, see Chapter 6 of Vadhan's monograph.

Gil Cohen

November 16, 2020

Overview

Random walks

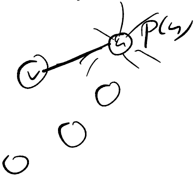
Let $G = (V, E)$ be an undirected graph, and $\mathbf{p}$ a probability distribution on V, thought of as a vector $\mathbf{p} \in \mathbb{R}^V$.

A random step on G, starting from a probability distribution $\mathbf{p}$, is the process in which we

1. Sample v according to $\mathbf{p}$;
2. Sample a neighbor u of v uniformly at random, and return u.

If $\mathbf{p}_{\text{new}}$ is distribution over V after taking a random step, then for every $v \in V$,

$$\mathbf{p}_{\text{new}}(v) = \sum_{u \in \Gamma(v)} \frac{\mathbf{p}(u)}{\deg(u)}.$$



Random walks

$$D_G = \begin{array}{cc} & u \\ u & \begin{pmatrix} & & 0 \\ 0 & \deg(u) & \\ & & \end{pmatrix} \end{array}$$

$$\mathbf{p}_{\text{new}}(v) = \sum_{u \in \Gamma(v)} \frac{\mathbf{p}(u)}{\deg(u)}.$$

Note that

$$\mathbf{p}_{\text{new}} = \mathbf{W}_G \mathbf{p} = \mathbf{M}_G \mathbf{D}_G^{-1} \mathbf{p}\,.$$

A length t random walk is the probabilistic process of taking t consecutive random steps. The corresponding distributions are given by

$$\mathbf{p}_t = \mathbf{W}\mathbf{p}_{t-1} = \mathbf{W}^2 p_{t-2} = \cdots = \mathbf{W}^t \mathbf{p}_0.$$

The normalized adjacency matrix

We define to the normalized adjacency matrix of G by $W = MD^{-1}$

$B = \sum \lambda_i \psi_i \psi_i^T \quad \lambda_i \geq 0$

$\sqrt{B} = \sum \sqrt{\lambda_i} \psi_i \psi_i^T$

$$\mathbf{A}_G = \mathbf{D}_G^{-1/2} \mathbf{M}_G \mathbf{D}_G^{-1/2}.$$

Note that $\mathbf{A}_G$ is symmetric for undirected graph G and that $\mathbf{A}_G = \mathbf{D}_G^{-1/2} \mathbf{W}_G \mathbf{D}_G^{1/2}$. $\quad W = D^{\frac{1}{2}} A D^{-\frac{1}{2}}$

Claim

ψ is an eigenvector of $\mathbf{A}$ of eigenvalue ω if and only if $\mathbf{D}^{1/2}\psi$ is an eigenvector of $\mathbf{W}$ of eigenvalue ω.

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

$$= D^{\frac{1}{2}} A \psi = \omega (D^{\frac{1}{2}}\psi)$$

$(D^{\frac{1}{2}}\psi)^T (D^{\frac{1}{2}}\psi')$

$\psi^T D \psi'$

The normalized adjacency matrix

A fact you should know (and prove to yourself!)

Lemma

For $n \times n$ matrices $\mathbf{A}, \mathbf{B}$,

$$\phi_{\mathbf{AB}}(x) = \phi_{\mathbf{BA}}(x).$$

More generally, if $\mathbf{A}$ is an $n \times m$ matrix and $\mathbf{B}$ an $m \times n$ matrix with $n > m$ then

$$\phi_{\mathbf{AB}}(x) = x^{n-m}\phi_{\mathbf{BA}}(x).$$

In particular, the spectrum remains the same (and the kernel increase when $n \neq m$).

The normalized adjacency matrix

We denote the eigenvalues of $\mathbf{W}$ by $\omega_1 \geq \omega_2 \geq \cdots \geq \omega_n$. Note that the degree vector $\mathbf{d}$ is an eigenvector of $\mathbf{W}$ of eigenvalue $\omega_1 = 1$. Indeed,

$d(v) = \deg(v)$

$$\mathbf{W}\mathbf{d} = (\mathbf{M}\mathbf{D}^{-1})d = \mathbf{M}(\mathbf{D}^{-1}\mathbf{d}) = \mathbf{M}\mathbf{1} = \mathbf{d}.$$

Define

$D^{-\frac{1}{2}} d = \sqrt{d}$

$$\psi_1 = \frac{\sqrt{\mathbf{d}}}{\|\sqrt{\mathbf{d}}\|} = \sqrt{\frac{\mathbf{d}}{\mathbf{1}^T\mathbf{d}}}.$$

Thus, ψ_1 is an eigenvector of $\mathbf{A}$ of eigenvalue 1. The Perron-Frobenius Theorem implies that $\operatorname{Spec}(\mathbf{W}) = \operatorname{Spec}(\mathbf{A}) \subset [-1, 1]$.

$\mu_1 \geq -\mu_n$

$\mu_1 > \mu_2$

$|\mu_n|$

The stable distribution

We denote $\omega(G) = \max(\omega_2, -\omega_n)$. By Perron-Frobenius, G is connected and not bipartite if and only if $\omega(G) < 1$.

Theorem

Assume that G is connected and not bipartite. Then, a random walk from any initial distribution converges to the stable distribution

$$\boldsymbol{\pi} = \frac{\mathbf{d}}{\mathbf{1}^T\mathbf{d}}.$$

Extra space for the proof

$$w_1 \quad \psi_1 = \sqrt{\frac{d}{1^T d}}$$

$$A = \sum_{i=1}^{n} \omega_i \psi_i \psi_i^T = 1 \cdot \psi_1 \psi_1^T + \sum_{i=2}^{n} \omega_i \psi_i \psi_i^T$$

$$W = D^{\frac{1}{2}} \psi_1 \psi_1^T D^{-\frac{1}{2}} + \sum_{i=2}^{n} \omega_i D^{\frac{1}{2}} \psi_i \psi_i^T D^{-\frac{1}{2}}$$

$\forall p$

t

$$W^t p = (D^{\frac{1}{2}} A D^{-\frac{1}{2}})^t p = D^{\frac{1}{2}} A^t D^{-\frac{1}{2}} p$$

$$W^t p = D^{\frac{1}{2}} \psi_1 \psi_1^T D^{-\frac{1}{2}} p + \sum_{i=2}^{n} \omega_i^t D^{\frac{1}{2}} \psi_i \psi_i^T D^{-\frac{1}{2}} p$$

$$\downarrow t \to \infty$$

$$0$$

$$\psi_1 = \sqrt{\frac{d}{1^T d}}$$

$$\pi = \frac{d}{1^T d} = \underbrace{D^{\frac{1}{2}} \sqrt{d}}_{d} \underbrace{\sqrt{d}^T D^{-\frac{1}{2}}}_{= 1^T} \frac{1}{1^T d} p$$

1

The rate of convergence

Theorem

Let $p_0 = e(u)$ for some $u \in V$. Then, for every $v \in V$,

$$|p_t(v) - \pi(v)| \leq \omega(G)^t \cdot \sqrt{\frac{\deg(v)}{\deg(u)}}.$$

$$W = D^{\frac{1}{2}} A D^{-\frac{1}{2}}$$

$$W^t = D^{\frac{1}{2}} A^t D^{-\frac{1}{2}}$$

$$W^t e(u) = \pi + \sum_{i=2}^{n} \omega_i^t D^{\frac{1}{2}} \psi_i \psi_i^T D^{-\frac{1}{2}} e(u)$$

Extra space for the proof

$$e(v)^T W^t e(u) = \pi(v) + \underbrace{\sum_{i=2}^{n} \omega_i^t e(v)^T D^{\frac{1}{2}} \psi_i \psi_i^T D^{-\frac{1}{2}} e(u)}_{\text{error term}}$$

$$e(v)^T W^t e(u) = P_t(v)$$

$$|P_t(v) - \pi(v)| \leq \sum_{i=2}^{n} |\omega_i|^t \left| \underbrace{e(v)^T D^{\frac{1}{2}}}\psi_i \psi_i^T \underbrace{D^{-\frac{1}{2}} e(u)} \right|$$

$$\|v\|^2 = \sum_{i=1}^{n} (v^T \psi_i)^2$$

$$\leq \omega^t \sqrt{\frac{\deg v}{\deg u}} \cdot \sum_{i=2}^{n} \left| e(v)^T \psi_i \psi_i^T e(u) \right|$$

$$\leq \omega^t \sqrt{\frac{\deg v}{\deg u}} \cdot \underbrace{\sqrt{\sum_{i=2}^{n} (e(v)^T \psi_i)^2}}_{\|e(v)\|} \underbrace{\sqrt{\sum_{i=2}^{n} (\psi_i^T e(u))^2}}_{\|e(u)\|}$$

Overview

Seeded extractors

Definition

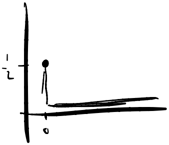
A distribution X has min entropy k if $\forall x, \Pr[X = x] \leq 2^{-k}$.

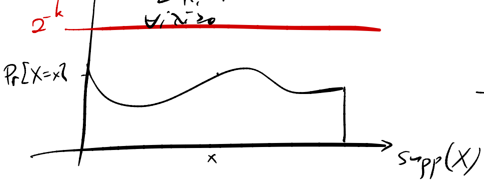
Claim

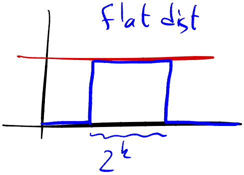
A distribution with min entropy k is a convex combination of distributions each is uniform over a set of size at least 2^k.

$$P = \sum_{i=1}^{m} \lambda_i P_i \qquad \lambda P_1 + (1-\lambda) P_2$$

$$\sum \lambda_i = 1 \qquad \forall i \; \lambda_i \geq 0$$







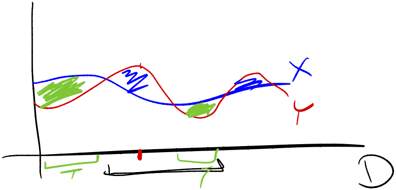
Seeded extractors

Definition

The statistical distance (aka total variation distance) between two distribution X, Y with support contained in D is given by

$$\mathbf{SD}(X, Y) = \max_{T \subseteq D} |\Pr[X \in T] - \Pr[Y \in T]|.$$

If $\mathbf{SD}(X, Y) \leq \varepsilon$ we write $X \approx_\varepsilon Y$.



bit-fixing

$\underbrace{01{*}01{*}{*}110}_{n}$

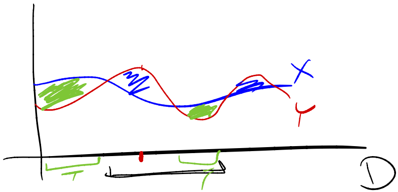
$**$ $\qquad \frac{n}{3}+1$

$\frac{1}{2}\log k$

Seeded extractors

Claim

$$\mathbf{SD}(X, Y) = \frac{1}{2} \cdot \|X - Y\|_1 = \frac{1}{2} \cdot \sum_{z \in D} |X(z) - Y(z)|.$$



Seeded extractors

Min entropy of X is k

Definition

A function $\text{Ext} : \{0,1\}^n \times \{0,1\}^s \to \{0,1\}^m$ is a (k,ε)-seeded extractor if for every k-source X, $\text{Ext}(X,Y) \approx_\varepsilon U_m$.

$E(x,y)$

Proposition

For every $n \geq k$ and ε there exists a (k,ε)-seeded extractor with

$$s = \log(n-k) \overset{+1}{} + 2\log\frac{1}{\varepsilon} + O(1)$$

$$m = k \underset{+s}{} - 2\log\frac{1}{\varepsilon} - O(1).$$

$m = 1$

Seeded extractors from random walks

The construction of Ext.

Set $s = td$. Consider a $D = 2^d$-regular graph G on $N = 2^n$ vertices. On input $x \in \{0,1\}^n$, $y \in \{0,1\}^s$ proceed as follows:

1. Interpret the given sample $x \sim X$ as a vertex.
2. Take a length-t random walk on G and return the last vertex on the path.

analysis.

While we can proceed as before, we will take a slightly different approach. Write $\mathbf{p}$ for the distribution induced by X.

Seeded extractors from random walks

$w^t p$ $\qquad \|p_t - \pi\|_2^2 = \sum_z \left(p_t(z) - \frac{1}{N}\right)^2$

Claim

It holds that $\|\mathbf{p}_t - \boldsymbol{\pi}\|_2 \leq \cancel{2}\underline{\omega(G)^t \cdot 2^{-k/2}}.$ $\quad \left(\omega(G)^t\left(2^{-\frac{k}{2}} + 2^{-\frac{n}{2}}\right)\right)$

Claim

For every $\mathbf{x} \in \mathbb{R}^N$, $\|x\|_1 \leq \sqrt{N} \cdot \|x\|_2.$

Hence,

$$\mathbf{SD}(\mathrm{Ext}(X,Y),U) \leq 2\underbrace{\omega(G)}^{\frac{1}{\sqrt{D}}}{}^t \cdot 2^{(n-k)/2}. \leq \varepsilon$$

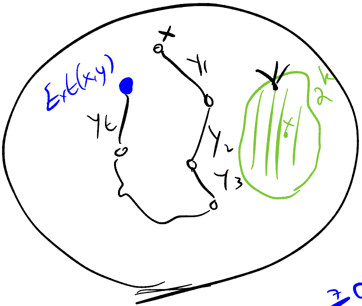
We will later see that there are graphs with $\omega(G) = O\left(\frac{1}{\sqrt{D}}\right)$. $\frac{2\sqrt{D}}{\sqrt{D}}$

Thus, $s = n - k + 2\log\frac{1}{\varepsilon} + O(1)$.

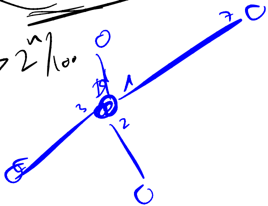
$s \geq n - k + 2\log\frac{1}{\varepsilon} + O(1)$ $\qquad \left(2D\right)^{t/2} \geq 2^{\frac{n-k}{2}} \cdot \frac{1}{\varepsilon}$

$2^{\frac{1+\log D}{2}} = 2^{s/2}$

Extra space for the proof



$$\underbrace{2^k \cdot D^t}_{} > 2^n/100$$



$D = 2^d$ regular

$N = 2^n \leftarrow$

$x \in \{0,1\}^n = V$

$y = \{0,1\}^s$

$s = dt$

$y = (y_1, y_2, \dots, y_t) \quad y_i \in \{0,1\}^d$

Extra space for the proof

$$\|p_t - \pi\|_2^2 = \|W^t p - \pi\|_2^2$$

$$\boxed{\begin{gathered}\|Ax\|_2^2 \\ \| \\ x^T A^T A x = x^T A^2 x\end{gathered}}$$

$$= \|W^t \underbrace{(p-\pi)}_{\pi \perp p-\pi}\|_2^2$$

$$= (p-\pi)^T (W^t)^T W^t (p-\pi)$$

$$= (p-\pi)^T W^{2t} (p-\pi) \le \omega(G)^{2t} (p-\pi)^T (p-\pi)$$

$$1 = \omega_1 \ge \omega_2 \ge \omega_3 \ge \cdots \ge \omega_n$$

Extra space for the proof

$\pi^T = \left(\frac{1}{N}, \dots, \frac{1}{N}\right)$

$$(p-\pi)^T(p-\pi) = p^T p \underbrace{-2\underbrace{p^T\pi}_{\frac{1}{N}} + \underbrace{\pi^T\pi}_{\frac{1}{N}}}_{-\frac{1}{N}}$$

$$p^T p = \sum_{i=1}^{N} p_i^2 \leq$$

$$\leq \underbrace{\max_i |p_i|}_{\leq 2^{-k}} \cdot \underbrace{\sum_{i=1}^{N} |p_i|}_{1}$$

$$\underbrace{\boxed{2^{-k} \dots 2^{-k}}}_{2^k} \underbrace{\boxed{0}}_{2^n - 2^k}$$

$$p^T\pi = \frac{p_1}{N} + \frac{p_2}{N} + \dots + \frac{p_N}{N} = \frac{1}{N}$$

$$= (2^{-k})^2 \cdot 2^k = 2^{-k}$$

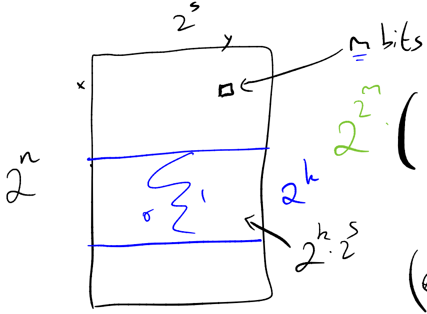
$$\|x\|_1 = \sum_{i=1}^{N} |x_i| \cdot 1 \leq \underbrace{\sqrt{\sum_i x_i^2}}_{\|x\|_2} \underbrace{\sqrt{\sum_i 1^2}}_{\sqrt{N}}$$

Extra space for the proof

$$\|p_t - \pi\|_2^2 \le \omega(G)^{2t} \underbrace{(p-\pi)^T(p-\pi)}_{\le 2^{-k}}$$

$$\|p_t - \pi\|_2 \le \omega(G)^t \, 2^{-k/2}$$

Extra space for the proof



$$\binom{n}{k} \le \left(\frac{ne}{k}\right)^k$$

$$2^{2^m} \binom{2^n}{2^k} e^{-\varepsilon^2 \cdot 2^{k+s}} < 1$$

$$\uparrow$$

$$\left(e\, 2^{n-k}\right)^{2^k} < e^{\varepsilon^2 2^{k+s}}$$

$$2^m + (n-k)\cancel{2^k} < \varepsilon^2 \cancel{2^{k}}{}^{+s}$$

$$s \ge \log(n-k) + 2\log\frac{1}{\varepsilon}$$