Expander and Derandomization

Many derandomization results are based on the assumption that certain random/hard objects exist.

Some unconditional derandomization can be achieved using explicit constructions of pseduorandom objects.

Synopsis

1. Basic Linear Algebra
2. Random Walk
3. Expander Graph
4. Explicit Construction of Expander Graph
5. Reingold's Theorem

Basic Linear Algebra

Three Views

All boldface lower case letters denotes column vectors.

Matrix = Linear transformation : $\mathbf{Q}^n \rightarrow \mathbf{Q}^m$

1. $f(\mathbf{u}+\mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$, $f(c\mathbf{u}) = cf(\mathbf{u})$
2. the matrix M_f corresponding to f has $f(\mathbf{e}_j)$ as the j-th column

---

Interpretation of $\mathbf{v} = A\mathbf{u}$

1. Dynamic view: $\mathbf{u}$ is transformed to $\mathbf{v}$, movement in one basis
2. Static view: $\mathbf{u}$ in the column basis is the same as $\mathbf{v}$ in the standard basis, movement of basis

---

Equation, Geometry (row picture), Algebra (column picture)

- Linear equation, hyperplane, linear combination

Suppose M is a matrix, $\mathbf{c}_1, \ldots, \mathbf{c}_n$ are column vectors, and $\mathbf{r}_1, \ldots, \mathbf{r}_n$ are row vectors.

$$\begin{aligned} M(\mathbf{c}_1, \ldots, \mathbf{c}_n) &= (M\mathbf{c}_1, \ldots, M\mathbf{c}_n) && (1) \\ (\mathbf{c}_1, \ldots, \mathbf{c}_n) \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_n \end{pmatrix} &= \mathbf{c}_1\mathbf{r}_1 + \mathbf{c}_2\mathbf{r}_2 + \ldots + \mathbf{c}_n\mathbf{r}_n && (2) \end{aligned}$$

Inner Product, Projection, Orthogonality

1. Inner product $\mathbf{u}^\dagger\mathbf{v}$ measures the degree of colinearity of $\mathbf{u}$ and $\mathbf{v}$
   - $\frac{\mathbf{u}^\dagger\mathbf{v}}{\|\mathbf{u}\|}\frac{\mathbf{u}}{\|\mathbf{u}\|}$ is the projection of $\mathbf{v}$ onto $\mathbf{u}$, where $\|\mathbf{u}\| = \sqrt{\mathbf{u}^\dagger\mathbf{u}}$ is the length of $\mathbf{u}$
   - $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if $\mathbf{u}^\dagger\mathbf{v} = 0$
   - $\frac{\mathbf{u}}{\|\mathbf{u}\|}$ is the normalization of $\mathbf{u}$
   - projection matrix $P = \frac{\mathbf{u}\mathbf{u}^\dagger}{\mathbf{u}^\dagger\mathbf{u}} = \frac{\mathbf{u}}{\|\mathbf{u}\|}\cdot\frac{\mathbf{u}^\dagger}{\|\mathbf{u}\|}$
   - suppose $\mathbf{u}_1, \ldots, \mathbf{u}_m$ are linearly independent. the projection of $\mathbf{v}$ onto the subspace spanned by $\mathbf{u}_1, \ldots, \mathbf{u}_m$ is $P\mathbf{v}$, where the projection matrix P is $A(A^\dagger A)^{-1}A^\dagger$.
     if $\mathbf{u}_1, \ldots, \mathbf{u}_m$ are orthonormal, $P = \mathbf{u}_1\mathbf{u}_1^\dagger + \ldots + \mathbf{u}_m\mathbf{u}_m^\dagger = I_m$.
2. Basis, orthonormal basis, orthogonal matrix
3. $Q^{-1} = Q^\dagger$ for every orthogonal matrix Q
   - Gram-Schmidt orthogonalization, $A = QR$

Cauchy-Schwartz Inequality. $\cos\theta = \frac{\mathbf{u}^\dagger\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} \leq 1$.

Fixpoints for Linear Transformation

We look for fixpoints of a linear transformation $A : \mathbf{R}^n \rightarrow \mathbf{R}^n$.

$$A\mathbf{v} = \lambda\mathbf{v}.$$

If there are n linear independent fixpoints $\mathbf{v}_1, \ldots, \mathbf{v}_n$, then every $\mathbf{v} \in \mathbf{R}^n$ is some linear combination $c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n$. By linearity,

$$A\mathbf{v} = c_1 A\mathbf{v}_1 + \ldots + c_n A\mathbf{v}_n = c_1\lambda_1\mathbf{v}_1 + \ldots + c_n\lambda_n\mathbf{v}_n.$$

If we think of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ as a basis, the effect of the transform A is to stretch the coordinates in the directions of the axes.

Eigenvalue, Eigenvector, Eigenmatrix

If $A - \lambda I$ is singular, an eigenvector $\mathbf{x}$ satisfies $\mathbf{x} \neq \mathbf{0}$, $A\mathbf{x} = \lambda\mathbf{x}$; and λ is the eigenvalue.

1. $S = [\mathbf{x_1}, \ldots, \mathbf{x_n}]$ is the eigenmatrix. By definition $AS = S\Lambda$.
2. If $\lambda_1, \ldots, \lambda_n$ are different, $\mathbf{x_1}, \ldots, \mathbf{x_n}$ are linearly independent.
3. If $\mathbf{x_1}, \ldots, \mathbf{x_n}$ are linearly independent, $A = S\Lambda S^{-1}$.

Suppose $c_1\mathbf{x}_1 + \ldots + c_n\mathbf{x}_n = 0$. Then $c_1\lambda_1\mathbf{x}_1 + \ldots + c_n\lambda_n\mathbf{x}_n = 0$. It follows that $c_1(\lambda_1 - \lambda_n)\mathbf{x}_1 + \ldots + c_{n-1}(\lambda_{n-1} - \lambda_n)\mathbf{x}_{n-1} = 0$. By induction we eventually get $c_1(\lambda_1 - \lambda_2)\ldots(\lambda_1 - \lambda_n)\mathbf{x}_1 = 0$. Thus $c_1 = 0$. Similarly $c_2 = \ldots = c_n = 0$.

- We shall write the spectrum $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$.
- $\rho(A) = |\lambda_1|$ is called spectral radius.

Similarity Transformation

Similarity Transformation = Change of Basis

1. A is similar to B if $A = MBM^{-1}$ for some invertible M.
2. $\mathbf{v}$ is an eigenvector of A iff $M^{-1}\mathbf{v}$ is an eigenvector of B.

A and B describe the same transformation using different bases.

1. The basis of B consists of the column vectors of M.
2. A vector $\mathbf{x}$ in the basis of A is transformed into the vector $M^{-1}\mathbf{x}$ in the basis of B, that is $\mathbf{x} = M(M^{-1}\mathbf{x})$.
3. B then transforms $M^{-1}\mathbf{x}$ into some $\mathbf{y}$ in the basis of B.
4. In the basis of A the vector $A\mathbf{x}$ is $M\mathbf{y}$.

Fact. Similar matrices have the same eigenvalues.

Triangularization

Diagonalization transformation is a special case of similarity transformation. In diagonalization Q provides an orthogonal basis.

Question. Is every matrix similar to a diagonal matrix?

Schur's Lemma. For each matrix A there is a unitary matrix U such that $T = U^{-1}AU$ is triangular. The eigenvalues of A appear in the diagonal of T.

Diagonalization

What are the matrices that are similar to diagonal matrices?

A matrix N is normal if $NN^\dagger = N^\dagger N$.

Theorem. A matrix N is normal iff $T = U^{-1}NU$ is diagonal iff N has a complete set of orthonormal eigenvectors.

Proof.

If N is normal, T is normal. It follows from $T^\dagger = T$ that T is diagonal. If T is diagonal, it is the eigenvalue matrix of N, and $NU = UT$ says that the column vectors of U are precisely the eigenvectors. □

Hermitian Matrix and Symmetric Matrix

	real matrix	**complex matrix**
length	$\|x\| = \sqrt{\sum_{i\in[n]} x_i^2}$	$\|x\| = \sqrt{\sum_{i\in[n]} \|x_i\|^2}$
conjugate transpose	$A^\dagger$	$A^\dagger$
inner product	$\mathbf{x}^\dagger\mathbf{y} = \sum_{i\in[n]} x_i y_i$	$\mathbf{x}^\dagger\mathbf{y} = \sum_{i\in[n]} \overline{x}_i y_i$
orthogonality	$\mathbf{x}^\dagger\mathbf{y} = 0$	$\mathbf{x}^\dagger\mathbf{y} = 0$
symmetric/Hermitian	$A^\dagger = A$	$A^\dagger = A$
diagonalization	$A = Q\Lambda Q^\dagger$	$A = U\Lambda U^\dagger$
orthogonal/unitary	$Q^\dagger Q = I$	$U^\dagger U = I$

Fact. If $A^\dagger = A$, then $\mathbf{x}^\dagger A\mathbf{x} = (\mathbf{x}^\dagger A\mathbf{x})^\dagger$ is real for all complex $\mathbf{x}$.

Fact. If $A^\dagger = A$, the eigenvalues are real since $\mathbf{v}^\dagger A\mathbf{v} = \lambda\mathbf{v}^\dagger\mathbf{v} = \lambda\|\mathbf{v}\|^2$.

Fact. If $A^\dagger = A$, the eigenvectors of different eigenvalues are orthogonal.

Fact. $\|U\mathbf{x}\|^2 = \|\mathbf{x}\|^2$ and $\|Q\mathbf{x}\|^2 = \|\mathbf{x}\|^2$.

Spectral Theorem

Theorem. Every Hermitian matrix A can be diagonalized by a unitary matrix U. Every symmetric matrix A can be diagonalized by an orthogonal matrix Q.

$$\begin{aligned} U^\dagger A U &= \Lambda, \\ Q^\dagger A Q &= \Lambda. \end{aligned}$$

The eigenvalues are in Λ; the orthonormal eigenvectors are in Q respectively U.

Corollary. Every Hermitian matrix A has a spectral decomposition.

$$A = U\Lambda U^\dagger \overset{(1)(2)}{=} \sum_{i\in[n]} \lambda_i \mathbf{u}_i \mathbf{u}_i^\dagger .$$

Notice that $I = UU^\dagger \overset{(2)}{=} \sum_{i\in[n]} \mathbf{u}_i \mathbf{u}_i^\dagger$.

Positive Definite Matrix

Symmetric matrixes with positive eigenvalues are at the center of many applications.

A symmetric matrix A is positive definite if $\mathbf{x}^\dagger A\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$.

Theorem. Suppose A is symmetric. The following are equivalent.

1. $\mathbf{x}^\dagger A\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$.
2. $\lambda_i > 0$ for all the eigenvalues λ_i.
3. $A = R^\dagger R$ for some matrix R with independent columns.

If we replace $>$ by $\geq$, we get positive semidefinite matrices.

Singular Value Decomposition

Consider an $m \times n$ matrix. Both $AA^\dagger$ and $A^\dagger A$ are symmetric.

---

1. $AA^\dagger$ is positive semidefinite since $\mathbf{x}^\dagger AA^\dagger \mathbf{x} = \|A^\dagger \mathbf{x}\|^2 \geq 0$.
2. $AA^\dagger = U\Sigma U^\dagger$, where U consists of the orthonormal eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_m$ and Σ is the diagonal matrix made up from the eigenvalues $\sigma_1^2 \geq \ldots \geq \sigma_r^2$.
3. $A^\dagger A = V\Sigma' V^\dagger$.
4. $AA^\dagger \mathbf{u}_i = \sigma_i^2 \mathbf{u}_i$ implies that $(\sigma_i^2, A^\dagger \mathbf{u}_i)$ is an eigenpair for $A^\dagger A$. So $\mathbf{v}_i = \frac{A^\dagger \mathbf{u}_i}{\|A^\dagger \mathbf{u}_i\|}$.
5. $\mathbf{u}_i^\dagger AA^\dagger \mathbf{u}_i = \mathbf{u}_i^\dagger \sigma_i^2 \mathbf{u}_i = \sigma_i^2$. So $\|A^\dagger \mathbf{u}_i\| = \sigma_i$.
6. $A\mathbf{v}_i = A\frac{A^\dagger \mathbf{u}_i}{\|A^\dagger \mathbf{u}_i\|} = \frac{\sigma_i^2 \mathbf{u}_i}{\sigma_i} = \sigma_i \mathbf{u}_i$.

Hence $AV = U\Sigma$.

Singular Value Decomposition

We call

1. $\sigma_1, \ldots, \sigma_r$ the singular values of A, and
2. $U\Sigma V^\dagger$ the singular value decomposition, or SVD, of A.

Lemma. If A is normal, then $\sigma_i = |\lambda_i|$ for all $i \in [n]$.

Proof.

Since A is normal, $A = U\Lambda U^\dagger$ by diagonalization. Now $A^\dagger A = AA^\dagger = U\Lambda^2 U^\dagger$. So the spectrum of $A^\dagger A/AA^\dagger$ is $\lambda_1^2, \ldots, \lambda_n^2$. □

Rayleigh Quotient

Suppose A is an $n \times n$ Hermitian matrix, $(\lambda_1, \mathbf{v}_1), \ldots, (\lambda_n, \mathbf{v}_n)$ are the eigenpairs.

The Rayleigh quotient of A and nonzero $\mathbf{x}$ is defined as follows:

$$R(A, \mathbf{x}) = \frac{\mathbf{x}^\dagger A \mathbf{x}}{\mathbf{x}^\dagger \mathbf{x}} = \frac{\sum_{i \in [n]} \lambda_i \|\mathbf{v}_i^\dagger \mathbf{x}\|^2}{\sum_{i \in [n]} \|\mathbf{v}_i^\dagger \mathbf{x}\|^2}. \tag{3}$$

It is clear from (3) that

- if $\lambda_1 \geq \ldots \geq \lambda_n$, then $\lambda_i = \max_{\mathbf{x} \perp \mathbf{v}_1, \ldots, \mathbf{x} \perp \mathbf{v}_{i-1}} R(A, \mathbf{x})$, and
- if $|\lambda_1| \geq \ldots \geq |\lambda_n|$, then $|\lambda_i| = \max_{\mathbf{x} \perp \mathbf{v}_1, \ldots, \mathbf{x} \perp \mathbf{v}_{i-1}} |R(A, \mathbf{x})|$.

One can use Rayleigh quotient to derive lower bound for λ_i.

Vector Norm

The norm of a vector is a measure of its magnitude/size/length.

A norm on $\mathbf{F}^n$ is a function $\|\_\| : \mathbf{F}^n \to \mathbf{R}^{\geq 0}$ satisfying the following:

1. $\|\mathbf{v}\| = 0$ iff $\mathbf{v} = \mathbf{0}$.
2. $\|a\mathbf{v}\| = |a| \cdot \|\mathbf{v}\|$.
3. $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$.

A vector space with a norm is called a normed vector space.

---

1. L^1-norm. $\|\mathbf{v}\|_1 = |\mathbf{v}_1| + \ldots + |\mathbf{v}_n|$.
2. L^2-norm. $\|\mathbf{v}\|_2 = \sqrt{|\mathbf{v}_1|^2 + \ldots + |\mathbf{v}_n|^2} = \sqrt{\mathbf{v}^\dagger \mathbf{v}}$.
3. L^p-norm. $\|\mathbf{v}\|_p = \sqrt[p]{|\mathbf{v}_1|^p + \ldots + |\mathbf{v}_n|^p}$.
4. L^∞-norm. $\|\mathbf{v}\|_\infty = \max\{|\mathbf{v}_1|, \ldots, |\mathbf{v}_n|\}$.

Matrix Norm

We define matrix norm in compatible with vector norm. Suppose $\mathbf{F}^n$ is a normed vector space over field $\mathbf{F}$.

An induced matrix norm is a function $\|\_\| : \mathbf{F}^{n\times n} \rightarrow \mathbf{R}^+$ satisfying the following properties.

1. $\|A\| = 0$ iff $A = \mathbf{0}$.
2. $\|aA\| = |a|\cdot\|A\|$.
3. $\|A + B\| \leq \|A\| + \|B\|$.
4. $\|AB\| \leq \|A\|\cdot\|B\|$.

Matrix Norm

A matrix norm measures the amplifying power of a matrix. Define

$$\|A\| = \max_{\mathbf{v}\neq\mathbf{0}} \frac{\|A\mathbf{v}\|}{\|\mathbf{v}\|}.$$

It satisfies (1-4). Additionally $\|A\mathbf{x}\| \leq \|A\|\cdot\|\mathbf{x}\|$ for all $\mathbf{x}$.

$$\begin{aligned} \|A\|_1 &= \max_{1\leq j\leq n} \sum_{i=1}^{n} |A_{i,j}|, \\ \|A\|_\infty &= \max_{1\leq i\leq n} \sum_{j=1}^{n} |A_{i,j}|. \end{aligned}$$

Lemma. $\rho(A) \leq \|A\|$.

Spectral Norm

$\|A\|_2$ is called the spectral norm of A.

$$\frac{1}{\sqrt{n}}\|A\|_1 \leq \|A\|_2 \leq \sqrt{n}\|A\|_1.$$

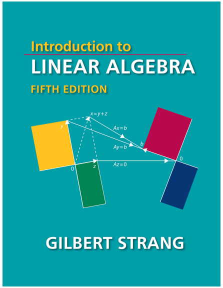
Lemma. $\|A\|_2 = \sigma_1$.

Corollary. If A is a normal matrix, then $\|A\|_2 = |\lambda_1|$.

Let $A^\dagger A = V\Sigma V^\dagger$, let $V = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$, and let $\mathbf{x} = a_1\mathbf{v}_1 + \ldots + a_n\mathbf{v}_n$. Then

$$\|A\mathbf{x}\|_2^2 = \mathbf{x}^\dagger(A^\dagger A\mathbf{x}) = \mathbf{x}^\dagger(\sum_{i\in[n]} \sigma_i^2 a_i \mathbf{v}_i) \leq \sigma_1^2\|\mathbf{x}\|_2^2.$$

The equality holds when $\mathbf{x} = \mathbf{v}_1$. Therefore $\|A\|_2 = \sigma_1$.



MIT Open Course

https://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010/video-lectures/

Random Walk

Graphs are the prime objects of study in combinatorics.

The matrix representation of graphs lends itself to an algebraic treatment to these combinatorial objects. It is especially effective in the treatment of regular graph.

Our digraph admit both self-loops and parallel edges. An undirected edge is seen as two directed edges in opposite directions.

In this lecture whenever we say graph, we mean undirected graph.

Random Walk Matrix

The reachability matrix M of a digraph G is defined by $M_{i,j} = 1$ if there is an edge from vertex j to vertex i; $M_{i,j} = 0$ otherwise.

The random walk matrix A of a d-regular digraph G is $\frac{1}{d}M$.

Let $\mathbf{p}$ be a probability distribution over the vertices of G and A is the random walk matrix of G. Then $A^k\mathbf{p}$ is the distribution after k-step random walk.

Random Walk Matrix

Consider the following periodic graph with dn vertices.

- The vertices are arranged in n layers, each consisting of d vertices. There is an edge from every vertex in the i-th layer to every vertex in the j-th layer, where $j = i + 1 \ mod \ n$.

Does $A^k\mathbf{p}$ converge to a stationary state?

Spectral Graph Theory

In spectral graph theory graph properties are characterized by graph spectrums.

Suppose G is a d-regular graph and A is the random walk matrix of G.

1. 1 is an eigenvalue of A and its associated eigenvector is the stationary distribution vector $\mathbf{1} = (\frac{1}{n}, \ldots, \frac{1}{n})^{\dagger}$. In other words $A\mathbf{1} = \mathbf{1}$.
2. All eigenvalues have absolute values ≤ 1.
3. G is disconnected if and only if 1 is an eigenvalue of multiplicity ≥ 2.
4. If G is connected, G is bipartite if and only if -1 is an eigenvalue of A.

---

2 and 3($\Leftarrow$) and 4($\Leftarrow$). Consider the entry with the largest absolute value.

Rate of Convergence

For a regular graph G with random walk matrix A, we define

$$\lambda_G \stackrel{\mathrm{def}}{=} \max_{\mathbf{p}} \frac{\|A\mathbf{p}-\mathbf{1}\|_2}{\|\mathbf{p}-\mathbf{1}\|_2} = \max_{\mathbf{v}\perp\mathbf{1}} \frac{\|A\mathbf{v}\|_2}{\|\mathbf{v}\|_2} = \max_{\mathbf{v}\perp\mathbf{1},\|\mathbf{v}\|_2=1} \|A\mathbf{v}\|_2,$$

where **p** is over all probability distribution vectors.

The two definitions are equivalent.

1. $(\mathbf{p}-\mathbf{1})\perp\mathbf{1}$ and $A\mathbf{p}-\mathbf{1}=A(\mathbf{p}-\mathbf{1})$.
2. For each $\mathbf{v}\perp\mathbf{1}$, $\mathbf{p}=\alpha\mathbf{v}+\mathbf{1}$ is a probability distribution for a sufficiently small α.

By definition $\|A\mathbf{v}\|_2 \le \lambda_G\|\mathbf{v}\|_2$ for all $\mathbf{v}$ such that $\mathbf{v}\perp\mathbf{1}$.

Lemma. $\lambda_G = |\lambda_2|$.

Let $\mathbf{v}_2, \ldots, \mathbf{v}_n$ be the eigenvectors corresponding to $\lambda_2, \ldots, \lambda_n$.

Given $\mathbf{x} \perp \mathbf{1}$, let $\mathbf{x} = c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n$. Then

$$\begin{array}{rcl} \|A\mathbf{x}\|_2^2 & = & \|\lambda_2 c_2 \mathbf{v}_2 + \ldots + \lambda_n c_n \mathbf{v}_n\|_2^2 \\ & = & \lambda_2^2 c_2^2 \|\mathbf{v}_2\|_2^2 + \ldots + \lambda_n^2 c_n^2 \|\mathbf{v}_n\|_2^2 \\ & \leq & \lambda_2^2 (c_2^2 \|\mathbf{v}_2\|_2^2 + \ldots + c_n^2 \|\mathbf{v}_n\|_2^2) \\ & = & \lambda_2^2 \|\mathbf{x}\|_2^2. \end{array}$$

So $\lambda_G^2 \leq \lambda_2^2$. The equality holds since $\|A\mathbf{v}_2\|_2^2 = \lambda_2^2 \|\mathbf{v}_2\|_2^2$.

The spectral gap γ_G of a graph G is defined by

$$\gamma_G = 1 - \lambda_G .$$

A graph G has spectral expansion γ, where $\gamma \in (0,1)$, if $\gamma_G \geq \gamma$.

In an expander the spectral expansion provides a bound on the expansion ratio.

Lemma. Let G be an n-vertex regular graph and $\mathbf{p}$ a probability distribution over the vertices of G. Then

$$\|A^{\ell}\mathbf{p} - \mathbf{1}\|_2 \leq \lambda_G^{\ell}\|\mathbf{p} - \mathbf{1}\|_2 < \lambda_G^{\ell}.$$

The first inequality holds because

$$\frac{\|A^{\ell}\mathbf{p} - \mathbf{1}\|_2}{\|\mathbf{p} - \mathbf{1}\|_2} = \frac{\|A^{\ell}\mathbf{p} - \mathbf{1}\|_2}{\|A^{\ell-1}\mathbf{p} - \mathbf{1}\|_2} \ldots \frac{\|A\mathbf{p} - \mathbf{1}\|_2}{\|\mathbf{p} - \mathbf{1}\|_2} \leq \lambda_G^{\ell}.$$

The second inequality holds because

$$\|\mathbf{p} - \mathbf{1}\|_2^2 = \|\mathbf{p}\|_2^2 + \|\mathbf{1}\|_2^2 - 2\langle \mathbf{p}, \mathbf{1}\rangle \leq 1 + \frac{1}{n} - 2\frac{1}{n} < 1.$$

In terms of random walk, λ_G bounds the speed of mixing time. [if G is bipartite, $\lambda_G = 1$.]

Lemma. If G is an n-vertex regular graph with self-loops at each vertex, $\gamma_G \geq \frac{1}{12n^2}$.

Let $\mathbf{u}$ be the unit vector such that $\mathbf{u} \perp \mathbf{1}$ and $\lambda_G = \|A\mathbf{u}\|_2$, and let $\mathbf{v} = A\mathbf{u}$.

- If we can prove $1 - \|\mathbf{v}\|_2^2 \geq \frac{1}{6n^2}$, we will get $\lambda_G = \|\mathbf{v}\|_2 \leq 1 - \frac{1}{12n^2}$, hence the lemma.
- It's easy to show $1 - \|\mathbf{v}\|_2^2 = \|\mathbf{u}\|_2^2 - \|\mathbf{v}\|_2^2 = \|\mathbf{u}\|_2^2 - 2\langle A\mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|_2^2 = \sum_{i,j} A_{i,j}(\mathbf{u}_i - \mathbf{v}_j)^2$.

Now $\mathbf{u}_i - \mathbf{u}_j \geq \frac{1}{\sqrt{n}}$ for some $i, j \in [n]$. Let $i \to i_1 \to \ldots \to i_k \to j$ be a shortest path. Then

$$1/\sqrt{n} \leq \mathbf{u}_i - \mathbf{u}_j \leq |\mathbf{u}_i - \mathbf{v}_i| + |\mathbf{v}_i - \mathbf{u}_{i_1}| + \ldots + |\mathbf{v}_{i_k} - \mathbf{u}_j| \tag{4}$$

$$\leq \sqrt{(\mathbf{u}_i - \mathbf{v}_i)^2 + (\mathbf{v}_i - \mathbf{u}_{i_1})^2 + \ldots + (\mathbf{v}_{i_k} - \mathbf{u}_j)^2} \cdot \sqrt{2D+1}, \tag{5}$$

where D is the diameter of G. Notice that there are at most $2D+1$ summands in (4). Thus

$$\sum_{i,j} A_{i,j}(\mathbf{u}_i - \mathbf{v}_j)^2 \geq 1/\left(dn(2D+1)\right) \geq 1/(6n^2)$$

by (5) and $A_{h,h}, A_{h,h+1} \geq 1/d$ and $D \leq 3n/(d+1)$. [see next slide.]

If two nodes, say u and v, in a shortest path between two nodes are of distance 3, then the neighbors of u plus u are disjoint from the neighbors of v plus v. It follows that

$$\frac{D}{3} \cdot (d+1) \leq n.$$

Randomized Algorithm for Undirected Connectivity

Corollary. Let G be a d-degree n-vertex graph with self-loop on every vertex. Let s, t be connected. Let $\ell > 24n^2 \log n$ and let X_ℓ denote the vertex distribution after ℓ step random walk from s. Then $\Pr[X_\ell = t] > \frac{1}{2n}$.

Graphs with self-loops are not bipartite. According to the Lemmas,

$$\|A^\ell \mathbf{e}_s - \mathbf{1}\|_2 < \left(1 - \frac{1}{12n^2}\right)^{24n^2 \log n} < \frac{1}{n^2}.$$

It follows that $\left(A^\ell \mathbf{e}_s\right)_t - \frac{1}{n} > -\frac{1}{n^2}$.

If the walk is repeated for $2n^2$ times, the error probability is reduced to below $\frac{1}{2^n}$.

Randomized Algorithm for Undirected Connectivity

Theorem. UPATH (Undirected Connectivity) is in **RL**.

An undirected graph can be turned into a non-bipartite regular graph by introducing enough self-loops.

Can the random algorithm for `UPATH` be derandomized? Recall that

$$\mathbf{L} \subseteq \mathbf{RL} \subseteq \mathbf{NL}.$$

Expander Graph

Expander graphs, defined by Pinsker in 1973, are sparse and well connected. They behave approximately like complete graphs.

- Sparsity should be understood in an asymptotic sense.

1. Fan Chung. Spectral Graph Theory. American Mathematical Society, 1997.
2. Hoory, Linial, and Wigderson. Expander Graphs and their Applications. Bulletin of the AMS, **43**, 439-561, 2006.

Well-connectedness can be characterized in a number of manners.

1. Algebraically, expanders are graphs whose second largest eigenvalue is bounded away from 1 by a constant.
2. Combinatorially, expanders are highly connected. Every set of vertices of an expander has a large boundary geometrically.
3. Probabilistically, expanders are graphs in which a random walk converges to the stationary distribution quickly.

Algebraic Property

Intuitively the faster random walk converges, the better the graph is connected.

According to Lemma, the smaller λ_G is, the faster random walk converges to $\mathbf{1}$.

Suppose $d \in \mathbf{N}$ and $\lambda \in (0,1)$ are constants.

A d-regular graph G with n vertices is an (n, d, λ)-graph if $\lambda_G \leq \lambda$.

$\{G_n\}_{n \in \mathbf{N}}$ is a (d, λ)-expander graph family if G_n is an (n, d, λ)-graph for all $n \in \mathbf{N}$.

Probabilistic Property

In an expander random walk converges to the uniform distribution in logarithmic steps.

$$\|A^{2\log_{\frac{1}{\lambda}}(n)}\mathbf{p}-\mathbf{1}\|_2 < \lambda^{2\log_{\frac{1}{\lambda}}(n)} = \frac{1}{n^2}.$$

In other words, the mixing time of an expander is logarithmic.

The mixing time of an n-vertex graph G is the minimal ℓ rendering true the inequality

$$\|A^{\ell}\mathbf{p}-\mathbf{1}\|_\infty < \frac{1}{2n}$$

for any vertex distribution, where A is the random walk matrix of G and $\mathbf{p}$.

The diameter of an n-vertex expander graph is $\Theta(\log n)$.

Combinatorial Property

Suppose $G = (V, E)$ is an n-vertex d-regular graph.

- Let $\overline{S}$ stand for $V \setminus S$ for $S \subseteq V$.
- Let $E(S, T)$ be the set of edges $i \rightarrow j$ with $i \in S$ and $j \in T$.
- Let $\partial S = E(S, \overline{S})$ for $|S| \leq \frac{n}{2}$.

The expansion constant h_G of G is defined as follows:

$$h_G = \min_{|S| \leq \frac{n}{2}} \frac{|\partial S|}{|S|}.$$

Constant $\rho > 0$. An n-vertex d-regular graph G is an (n, d, ρ)-edge expander if $\frac{h_G}{d} \geq \rho$.

- the number of edges leaving $S \geq \rho$ (the number of edges emitting from the nodes of S)

Existence of Expander

Theorem. Let $\epsilon > 0$. There exists $d = d(\epsilon)$ and $N \in \mathbf{N}$ such that for every $n > N$ there exists an $(n, d, \frac{1}{2} - \epsilon)$ edge expander.

Expansion and Spectral Gap

Theorem. Let $G = (V, E)$ be a finite, connected, d-regular graph. Then

$$\frac{\gamma_G}{2} \leq \frac{h_G}{d} \leq \sqrt{2(1-\gamma_G)}.$$

---

1. J. Dodziuk. Difference Equations, Isoperimetric Inequality and Transience of Certain Random Walks. Trans. AMS, 1984.
2. N. Alon and V. Milman. λ_1, Isoperimetric Inequalities for Graphs, and Superconcentrators. J. Comb. Theory, 1985.
3. N. Alon. Eigenvalues and Expanders. Combinatorica, 1986.

$d\cdot\frac{1-\lambda_G}{2} \leq h_G$

Let S be such that $|S| \leq \frac{n}{2}$ and $\frac{|\partial(S)|}{|S|} = h_G$. Define $\mathbf{x}\perp\mathbf{1}$ by $\mathbf{x}_i = \begin{cases} |\overline{S}|, & i \in S, \\ -|S|, & i \in \overline{S}. \end{cases}$

$$\begin{array}{rcl} \|\mathbf{x}\|_2^2 & = & n|S||\overline{S}|, \\ \mathbf{x}^\dagger A\mathbf{x} & = & (|\overline{S}|\mathbf{1}_S - |S|\mathbf{1}_{\overline{S}})^\dagger A(|\overline{S}|\mathbf{1}_S - |S|\mathbf{1}_{\overline{S}}) \\ & = & \frac{1}{d}\left(|\overline{S}|^2|E(S,S)| + |S|^2|E(\overline{S},\overline{S})| - 2|S||\overline{S}||E(S,\overline{S})|\right) \\ & = & \frac{1}{d}\left(dn|S||\overline{S}| - n^2|E(S,\overline{S})|\right), \end{array}$$

where $=$ is due to $d|S| = |E(S,\overline{S})| + |E(S,S)|$ and $d|\overline{S}| = |E(\overline{S},S)| + |E(\overline{S},\overline{S})|$.

The Rayleigh quotient $R(A,\mathbf{x})$ provides a lower bound for λ_G.

$$\lambda_G \geq \frac{\mathbf{x}^\dagger A\mathbf{x}}{\|\mathbf{x}\|_2^2} = \frac{1}{d}\frac{dn|S||\overline{S}| - n^2|E(S,\overline{S})|}{n|S||\overline{S}|} = 1 - \frac{1}{d}\cdot\frac{n}{|\overline{S}|}\cdot\frac{|\partial(S)|}{|S|} \geq 1 - \frac{2h_G}{d}.$$

$h_G \leq d \cdot \sqrt{2(1-\lambda_2)}$

Let $\mathbf{u} \perp \mathbf{1}$ be such that $A\mathbf{u} = \lambda_2 \mathbf{u}$. Write $\mathbf{u} = \mathbf{v} + \mathbf{w}$, where $\mathbf{v}$ respectively $\mathbf{w}$ is defined from $\mathbf{u}$ by replacing the negative respectively positive components by 0.

Wlog, assume that the number of positive components of $\mathbf{v}$ is $\leq \frac{n}{2}$.

Wlog, let the coordinates of $\mathbf{v}$ satisfy $\mathbf{v}_1 \geq \mathbf{v}_2 \geq \ldots \geq \mathbf{v}_n$. Then

$$\begin{aligned} \sum_{i,j} A_{i,j} |\mathbf{v}_i^2 - \mathbf{v}_j^2| &= 2 \sum_{i<j} A_{i,j} \sum_{k=i}^{j-1} (\mathbf{v}_k^2 - \mathbf{v}_{k+1}^2) = \frac{2}{d} \sum_{k=1}^{n/2} |\partial[k]| (\mathbf{v}_k^2 - \mathbf{v}_{k+1}^2) \\ &\geq \frac{2}{d} \sum_{k=1}^{n/2} h_G k (\mathbf{v}_k^2 - \mathbf{v}_{k+1}^2) = \frac{2h_G}{d} \|\mathbf{v}\|_2^2, \end{aligned} \tag{6}$$

where $=$ by swapping the two $\sum$'s. Notice that (i) $\mathbf{v}_k = 0$ for all $k > n/2$, and (ii) for each fixed k the term $\mathbf{v}_k^2 - \mathbf{v}_{k+1}^2$ appears once for every edge $i \to j$ such that $i \leq k < j$.

$h_G \leq d \cdot \sqrt{2(1-\lambda_2)}$

$\langle A\mathbf{v}, \mathbf{v}\rangle \geq \langle A\mathbf{v}, \mathbf{v}\rangle + \langle A\mathbf{w}, \mathbf{v}\rangle = \lambda_2 \|\mathbf{v}\|_2^2$ because $A\mathbf{u} = \lambda_2 \mathbf{u}$, $\langle \mathbf{w}, \mathbf{v}\rangle = 0$ and $\langle A\mathbf{w}, \mathbf{v}\rangle \leq 0$.

$$1 - \lambda_2 \geq 1 - \frac{\langle A\mathbf{v}, \mathbf{v}\rangle}{\|\mathbf{v}\|_2^2} = \frac{\|\mathbf{v}\|_2^2 - \langle A\mathbf{v}, \mathbf{v}\rangle}{\|\mathbf{v}\|_2^2} = \frac{\sum_{i,j} A_{i,j}(\mathbf{v}_i - \mathbf{v}_j)^2}{2\|\mathbf{v}\|_2^2}. \tag{7}$$

Using Cauchy-Schwartz Inequality,

$$\sum_{i,j} A_{i,j}(\mathbf{v}_i - \mathbf{v}_j)^2 \cdot \sum_{i,j} A_{i,j}(\mathbf{v}_i + \mathbf{v}_j)^2 \geq \left(\sum_{i,j} A_{i,j}|\mathbf{v}_i^2 - \mathbf{v}_j^2|\right)^2. \tag{8}$$

Now $\langle A\mathbf{v}, \mathbf{v}\rangle \leq \lambda_1 \|\mathbf{v}\|_2^2 = \|\mathbf{v}\|_2^2$. Therefore

$$2\|\mathbf{v}\|_2^2 \cdot \sum_{i,j} A_{i,j}(\mathbf{v}_i + \mathbf{v}_j)^2 \leq 2\|\mathbf{v}\|_2^2 \cdot (2\|\mathbf{v}\|_2^2 + 2\langle A\mathbf{v}, \mathbf{v}\rangle) \leq 8\|\mathbf{v}\|_2^4. \tag{9}$$

(6)+(7)+(8)+(9) implies $\sqrt{8(1-\lambda_2)} \geq \frac{2h_G}{d}$.

Combinatorial definition and algebraic definition are equivalent.

---

1. The inequality $d \cdot \frac{1-\lambda_G}{2} \leq h_G$ implies that if G is an (n, d, λ)-expander graph, then it is an $(n, d, \frac{1-\lambda}{2})$ edge expander.
2. The inequality $h_G \leq d \cdot \sqrt{2(1-\lambda_2)}$ implies that if G is an (n, d, ρ) edge expander, then it is an $(n, d, 1-\frac{\rho^2}{2})$-expander graph.

Convergence in Entropy

Rényi 2-Entropy:

$$H_2(\mathbf{p}) = \log\left(\frac{1}{\|\mathbf{p}\|_2^2}\right).$$

Fact. If A is the random walk matrix of an (n, d, λ)-expander, then $H_2(A\mathbf{p}) \geq H_2(\mathbf{p})$. The equality holds if and only if $\mathbf{p}$ is uniform.

Proof.

Let $\mathbf{p} = \mathbf{1} + \mathbf{w}$ such that $\mathbf{w} \perp \mathbf{1}$. Then $\langle A\mathbf{w}, \mathbf{1}\rangle = \mathbf{w}^\dagger A^\dagger \mathbf{1} = \mathbf{w}^\dagger A \mathbf{1} = \mathbf{w}^\dagger \mathbf{1} = 0$. Therefore

$$\|A\mathbf{p}\|_2^2 = \|\mathbf{1}\|_2^2 + \|A\mathbf{w}\|_2^2 \leq \|\mathbf{1}\|_2^2 + \lambda\|\mathbf{w}\|_2^2 \leq \|\mathbf{1}\|_2^2 + \|\mathbf{w}\|_2^2 = \|\mathbf{p}\|_2^2.$$

The equality holds when $\mathbf{p} = \mathbf{1}$. □

Random walks do not decrease randomness.

The smaller the spectral gap, or the larger the spectral expansion, the more expander graphs behave like random graphs. This is what the next lemma says.

Expander Mixing Lemma

Lemma. Let $G = (V, E)$ be an (n, d, λ)-expander graph. Let $S, T \subseteq V$. Then

$$\left| |E(S,T)| - \frac{d}{n}|S||T| \right| \leq \lambda d \sqrt{|S||T|}. \tag{10}$$

Notice that (10) implies

$$\left| \frac{|E(S,T)|}{dn} - \frac{|S|}{n} \cdot \frac{|T|}{n} \right| \leq \lambda. \tag{11}$$

The edge density ≈ the product of the vertex densities. This property is called mixing.

---

1. N. Alon and F. Chung. Explicit Construction of Linear Sized Tolerant Networks. Discrete Mathematics, 1988.

Proof of Expander Mixing Lemma

Let $[\mathbf{v}_1, \ldots, \mathbf{v}_n]$ be the eigenmatrix of G and set $\mathbf{v}_1 = (\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}})^\dagger$.

Let $\mathbf{1}_S = \sum_i \alpha_i \mathbf{v}_i$ and $\mathbf{1}_T = \sum_j \beta_j \mathbf{v}_j$ be the characteristic vectors of S, T respectively.

$$|E(S,T)| = (\mathbf{1}_S)^\dagger (dA) \mathbf{1}_T = \left(\sum_i \alpha_i \mathbf{v}_i\right)^\dagger (dA) \left(\sum_j \beta_j \mathbf{v}_j\right) = \sum_i d\lambda_i \alpha_i \beta_i.$$

Since $\alpha_1 = (\mathbf{1}_S)^\dagger \mathbf{v}_1 = \frac{|S|}{\sqrt{n}}$ and $\beta_1 = (\mathbf{1}_T)^\dagger \mathbf{v}_1 = \frac{|T|}{\sqrt{n}}$, by Cauchy-Schwartz Inequality,

$$\left| |E(S,T)| - \frac{d}{n}|S||T| \right| = \sum_{i=2}^{n} d\lambda_i \alpha_i \beta_i \leq d\lambda \sum_{i=2}^{n} \alpha_i \beta_i \leq d\lambda \|\alpha\|_2 \|\beta\|_2.$$

Finally observe that $\|\alpha\|_2 \|\beta\|_2 = \|\mathbf{1}_S\|_2 \|\mathbf{1}_T\|_2 = \sqrt{|S||T|}$.

Error Reduction for Random Algorithm

Suppose $A(x,r)$ is a random algorithm with error probability $1/3$. The algorithm uses $r(n)$ random bits on input x with $|x|=n$.

1. Reduce the error probability exponentially by repeating the algorithm $t(n)$ times.
2. Altogether $r(n)t(n)$ random bits are used.

The goal is to achieve the same error reduction rate using far fewer random bits (in fact $r(n)+O(t(n))$ random bits).

The key observation is that a t-step random walk in an expander graph looks like t vertices sampled uniformly and independently.

- Confer the inequality (11).

K_n is perfect from the viewpoint of random walk.

- No matter what distribution it starts with, random walk reaches the uniform distribution in one step.

Let $J_n = [\mathbf{1}, \ldots, \mathbf{1}]$ be the random walk matrix of K_n with self-loop.

Decomposition for Random Walk on Expander

Lemma. Suppose G is an (n, d, λ)-expander and A is its random walk matrix. Then $A = \gamma J_n + \lambda E$ for some E such that $\|E\| \leq 1$.

We may think of a random walk on an expander as a convex combination of two random walks of different type:

- with probability γ it walks randomly on a complete graph, and
- with probability λ it walks randomly according to an error matrix that does not amplify the distance to the uniform distribution.

Decomposition for Random Walk on Expander

We need to prove that $\|E\mathbf{v}\|_2 \leq \|\mathbf{v}\|_2$ for all $\mathbf{v}$, where E is defined by

$$E = \frac{1}{\lambda}(A - (1-\lambda)J_n).$$

The following proof methodology should now be familiar.

- Let $\alpha = \sum_{i\in[n]} \mathbf{v}_i$. Then $\mathbf{v} = \alpha\mathbf{1} + \mathbf{w}$ with $\mathbf{w}\perp\mathbf{1}$.
- $A\mathbf{1} = \mathbf{1}$ and $J_n\mathbf{1} = \mathbf{1}$. Consequently $E(\alpha\mathbf{1}) = \alpha\mathbf{1}$.
- $J_n\mathbf{w} = \mathbf{0}$ and $A\mathbf{w}\perp\alpha\mathbf{1}$. Hence $E\mathbf{w} = \frac{1}{\lambda}A\mathbf{w}$.
- $\|A\mathbf{w}\|_2 \leq \lambda\|\mathbf{w}\|_2$.

Thus $\|E\mathbf{v}\|_2^2 = \|\alpha\mathbf{1} + \frac{1}{\lambda}A\mathbf{w}\|_2^2 = \|\alpha\mathbf{1}\|_2^2 + \|\frac{1}{\lambda}A\mathbf{w}\|_2^2 \leq \|\alpha\mathbf{1}\|_2^2 + \|\mathbf{w}\|_2^2 = \|\mathbf{v}\|_2^2$. Done.

Expander Random Walk Theorem

Theorem. Let G be an (n,d,λ) expander graph, and let $B \subseteq [n]$ satisfy $|B| \leq \beta n$ for some $\beta \in (0,1)$. Let X_1 be a random variable denoting the uniform distribution on $[n]$ and let X_k be a random variable denoting a $k-1$ step random walk from X_1. Then

$$\Pr\left[\bigwedge_{i\in[k]} X_i \in B\right] \leq \left(\gamma\sqrt{\beta}+\lambda\right)^{k-1}.$$

Expander Random Walk Theorem

Let B_i stand for $X_i \in B$. We need to bound the following.

$$\Pr\left[\bigwedge_{i\in[k]} X_i \in B\right] = \Pr[B_1]\cdot\Pr[B_2|B_1]\ldots\Pr[B_k|B_1\ldots B_{k-1}]. \tag{12}$$

By seeing B as a diagonal matrix, we define the distribution vector $\mathbf{p}_i$ by

$$\mathbf{p}_i = \frac{BA}{\Pr[B_i|B_1\ldots B_{i-1}]}\cdot\ldots\cdot\frac{BA}{\Pr[B_2|B_1]}\cdot\frac{B\mathbf{1}}{\Pr[B_1]},$$

where $\frac{BA}{\Pr[B_2|B_1]}\cdot\frac{B\mathbf{1}}{\Pr[B_1]}$ for example is the normalization of $BA\cdot\frac{B\mathbf{1}}{\Pr[B_1]}$. So the probability in (12) is bounded by $\|(BA)^{k-1}B\mathbf{1}\|_1$. We will prove

$$\|(BA)^{k-1}B\mathbf{1}\|_2 \leq \frac{1}{\sqrt{n}}\left((1-\lambda)\sqrt{\beta}+\lambda\right)^{k-1}.$$

Expander Random Walk Theorem

Using Lemma,

$$\begin{aligned}\|BA\| &= \|B((1-\lambda)J_n+\lambda E)\| \ \leq\ (1-\lambda)\|BJ_n\|+\lambda\|BE\| \ =\ (1-\lambda)\sqrt{\beta}+\lambda\|BE\| \\ &\leq (1-\lambda)\sqrt{\beta}+\lambda\|B\|\|E\| \ \leq\ (1-\lambda)\sqrt{\beta}+\lambda.\end{aligned}$$

Therefore

$$\|(BA)^{k-1}B\mathbf{1}\|_2 \leq \frac{\sqrt{\beta}}{\sqrt{n}}\left((1-\lambda)\sqrt{\beta}+\lambda\right)^{k-1} \leq \frac{1}{\sqrt{n}}\left((1-\lambda)\sqrt{\beta}+\lambda\right)^{k-1}.$$

Suppose $\|\mathbf{v}\|_2 = 1$ and $\alpha = \sum_{i\in[n]} \mathbf{v}_i$. Then $\mathbf{v} = \alpha\mathbf{1}+\mathbf{w}$ and $\mathbf{w}\perp\mathbf{1}$. Now

- $\|BJ_n\mathbf{v}\|_2 = \|BJ_n\alpha\mathbf{1}\|_2 = \alpha\|B\mathbf{1}\|_2 \leq \sqrt{n}\|B\mathbf{1}\|_2 = \sqrt{n}\cdot\frac{\sqrt{\beta}}{\sqrt{n}} = \sqrt{\beta}$, and consequently
- $\|BJ_n\| = \max\{\|BJ_n\mathbf{v}\|_2 \mid \|\mathbf{v}\|_2 = 1\} = \sqrt{\beta}$.

Error Reduction for **RP**

Suppose $A(x,r)$ is a random algorithm with error probability β. Given input x with $n=|x|$, let $k=|r(n)|$.

Choose an explicit $(2^k,d,\lambda)$-graph $G=(V,E)$ with $V=\{0,1\}^k$.

Algorithm A'.

1. Pick $v_0 \in_{\mathrm{R}} V$.
2. Generate a random walk $v_0,\ldots,v_t$.
3. Output $\bigvee_{i=0}^{t} A(x,v_i)$.

By the Theorem, the error probability of A' is no more than $\left(\gamma\sqrt{\beta}+\lambda\right)^{t-1}$.

- Here B is the set of r's for which A errs on x.

Error Reduction for **BPP**

Algorithm A''.

1. Pick $v_0 \in_{\mathrm{R}} V$.
2. Generate a random walk $v_0, \ldots, v_t$.
3. Output $Maj\{A(x, v_i)\}_{i \in [t]}$.

Let $K \subseteq [t]$ be the set of samples for which A errs and $|K| \geq \frac{t+1}{2}$.

$$\Pr[\forall i \in K. v_i \in B] \leq \left(\gamma\sqrt{\beta} + \lambda\right)^{\frac{t-1}{2}} \leq \left(\frac{1}{4}\right)^{t-1},$$

assuming $\gamma\sqrt{\beta} + \lambda \leq 1/16$. By union bound,

$$\Pr[A'' \text{ fails}] \leq 2^t \left(\frac{1}{4}\right)^{t-1} = O(2^{-t}).$$

Explicit Construction of Expander Graph

Explicit Construction

In some applications expander graphs are small.

- An expander family $\{G_n\}_{n\in\mathbf{N}}$ is mildly explicit if there is a P-time algorithm that outputs the random walk matrix of G_n whenever the input is 1^n.

In some other applications expander graphs are large.

- An expander family $\{G_n\}_{n\in\mathbf{N}}$ is strongly explicit if there is a P-time algorithm that on input $\langle n, v, i\rangle$ outputs the index of the i-th neighbor of v.

We will look at several graph product operations. We then show how to use these operations to construct explicit expander graphs.

1. O. Reingold, S. Vadhan, and A. Wigderson. Entropy Waves, the Zig-Zag Graph Product, and New Constant-Degree Expanders and Extractors. FOCS, 2000.

Path Product

Suppose G, G' are n-vertex graphs with degree d respectively d'. Let A, A' be their random walk matrices.

The path product $G'G$ is defined by the random walk matrix $A'A$.

- $G'G$ is n-vertex dd'-degree.

Lemma. $\lambda_{G'G} \leq \lambda_{G'}\lambda_G$.

Proof.
$\lambda_{G'G} = \max_{\mathbf{v}\perp\mathbf{1}} \frac{\|A'A\mathbf{v}\|_2}{\|\mathbf{v}\|_2} = \max_{\mathbf{v}\perp\mathbf{1}} \frac{\|A'A\mathbf{v}\|_2}{\|A\mathbf{v}\|_2} \cdot \frac{\|A\mathbf{v}\|_2}{\|\mathbf{v}\|_2} \leq \max_{\mathbf{v}\perp\mathbf{1}} \frac{\|A'A\mathbf{v}\|_2}{\|A\mathbf{v}\|_2} \cdot \max_{\mathbf{v}\perp\mathbf{1}} \frac{\|A\mathbf{v}\|_2}{\|\mathbf{v}\|_2} \leq \lambda_{G'}\lambda_G$ using the fact that $A\mathbf{v}\perp\mathbf{1}$ whenever $\mathbf{v}\perp\mathbf{1}$. □

Lemma. $\lambda_{G^k} = (\lambda_G)^k$.

Proof.
$(\lambda_G)^k$ is the second largest eigenvalue of G^k. □

Tensor Product

Suppose G is an n-vertex d-degree graph and G' is an n'-vertex d'-degree graph.

The random walk matrix of the tensor product $G \otimes G'$ is nn'-vertex dd'-degree.

$$A \otimes A' = \begin{pmatrix} a_{11}A' & a_{12}A' & \cdots & a_{1n}A' \\ a_{21}A' & a_{22}A' & \cdots & a_{2n}A' \\ \vdots & \vdots & \ldots & \vdots \\ a_{n1}A' & a_{n2}A' & \cdots & a_{nn}A' \end{pmatrix}.$$

$(u, u') \to (v, v')$ in $G \otimes G'$ iff $u \to v$ in G and $u' \to v'$ in G'.

Tensor Product

Lemma. $\lambda_{G \otimes G'} = \max\{\lambda_G, \lambda_{G'}\}$.

If $(\lambda, \mathbf{v})$ is an eigenvpair of A and $(\lambda', \mathbf{v}')$ is an eigenpair of A', then $(\lambda\lambda', \mathbf{v} \otimes \mathbf{v}')$ is an eigenpair of $A \otimes A'$.

Rotation Matrix

Let A be the random walk matrix of an n-vertex regular graph G of degree D.

The rotation matrix $\widehat{A}$ is an $(nD)\times(nD)$ adjacent matrix such that $\widehat{A}_{(v,m),(u,l)} = 1$ if

- v is the i-th neighbor of u, and u is the j-th neighbor of v.

We need a graph here.

Zig-Zag Product

The zig-zag product $G \textcircled{z} H$ is the d^2-degree graph defined by $(I_n \otimes J_D)\widehat{A}(I_n \otimes J_D)$.

- G is an n-vertex regular graph of degree D, and A is the random walk matrix of G. H is a D-vertex regular graph of degree d.

---

picture here

(v, m) is the (i, j)-th neighbor of (u, l): l' is the i-th neighbor of l in H; v is the l'-th neighbor of u and u is the m'-th neighbor of v; m is the j-th neighbor of m' in H.

Zig-Zag Product

Lemma. $(I_n \otimes J_D)\widehat{A}(I_n \otimes J_D) = A \otimes J_D$.

$$\left((I_n \otimes J_D)\widehat{A}(I_n \otimes J_D)\right)_{(v,m),(u,l)} = \frac{1}{D} \cdot 1 \cdot \frac{1}{D} = \frac{1}{D} \cdot \frac{1}{D} = (A \otimes J_D)_{(v,m),(u,l)}.$$

Claim. If $\|C\|_2 \leq 1$ then $\lambda_C \leq 1$.

Proof.

$\lambda_C = \max_{\mathbf{v}\perp\mathbf{1}} \frac{\|C\mathbf{v}\|_2}{\|\mathbf{v}\|_2} \leq \max_{\mathbf{v}\perp\mathbf{1}} \frac{\|C\|_2\|\mathbf{v}\|_2}{\|\mathbf{v}\|_2} \leq \|C\|_2 \leq 1.$ □

Claim. $\lambda_{A+B} \leq \lambda_A + \lambda_B$ for symmetric matrices A, B.

Proof.

$\lambda_{A+B} = \max_{\mathbf{v}\perp\mathbf{1}} \frac{\|(A+B)\mathbf{v}\|_2}{\|\mathbf{v}\|_2} \leq \max_{\mathbf{v}\perp\mathbf{1}} \frac{\|A\mathbf{v}\|_2+\|B\mathbf{v}\|_2}{\|\mathbf{v}\|_2} \leq \lambda_A + \lambda_B.$ □

Zig-Zag Product

Lemma. $\lambda_{G \textcircled{z} H} \leq \lambda_G + 2\lambda_H$ and $\gamma_{G \textcircled{z} H} \geq \gamma_G \gamma_H^2$.

Let A, B and M be the random walk matrices of G, H and $G \textcircled{z} H$ respectively.

- $\widehat{A}$ is the $(nD)\times(nD)$ rotation matrix of G.
- $B = (1-\lambda_H)J_D + \lambda_H E$ for some E with $\|E\|_2 \leq 1$. This is the Lemma.

Now

$$\begin{aligned} M &= (I_n \otimes B)\widehat{A}(I_n \otimes B) = ((1-\lambda_H)I_n \otimes J_D + \lambda_H I_n \otimes E)\ \widehat{A}\ ((1-\lambda_H)I_n \otimes J_D + \lambda_H I_n \otimes E) \\ &= (1-\lambda_H)^2 (I_n \otimes J_D)\widehat{A}(I_n \otimes J_D) + \ldots = (1-\lambda_H)^2 (A \otimes J_D) + \ldots, \end{aligned}$$

where $=$ is due to Lemma. Using Lemma and the Claims, one gets

$$\lambda_M \leq (1-\lambda_H)^2 \lambda_{A \otimes J_D} + 1 - (1-\lambda_H)^2 \leq \max\{\lambda_G, \lambda_{J_D}\} + 2\lambda_H = \lambda_G + 2\lambda_H.$$

Comment on Zig-Zag Product

1. Typically $d \ll D$.
2. A t-step random walk uses $O(t \log d)$ rather than $O(t \log D)$ random bits.
3. The last lemma is useful when both λ_G and λ_H are small. If not, a different upper bound can be derived. Both upper bounds are discussed in the following paper.

[1] O. Reingold, S. Vadhan, and A. Wigderson. Entropy Waves, the Zig-Zag Graph Product, and New Constant Degree Expanders and Extractors. FOCS, 2000.

Expander Construction I

The crucial point of the zig-zag construction is that we can use a constant graph to build a constant degree graph family.

Let H be a $(D^4, D, 1/8)$-graph constructed by brute force. Define

$$\begin{array}{rcl} G_1 & = & H^2, \\ G_{k+1} & = & G_k^2 \textcircled{z} H. \end{array}$$

Fact. G_k is a $(D^{4k}, D^2, 1/2)$-graph.

Proof.

The base case is clear from Lemma, and the induction step is taken care of by the previous lemma. □

Expander Construction I

The time to access to a neighbor of a vertex is given by the following recursive equation

$$\begin{aligned}
\mathrm{time}(G_{k+1}) &= 2 \cdot \mathrm{time}(G_k) + \mathrm{time}(H) \\
&= 2^k \cdot \mathrm{time}(H^2) + (2^{k-1} + \ldots + 2 + 1) \cdot \mathrm{time}(H) \\
&= 2^{O(k)} \\
&= \mathrm{poly}(|G_{k+1}|).
\end{aligned}$$

The expander family is mildly explicit but not strongly explicit.

Comment on Expander Construction I

Both the size of the graph and the time to compute a neighbor grow exponentially. This suggests to use tensor product to expand the size of graph double exponentially.

	Size	Degree	Expansion
Path Product	–	$\uparrow$	$\Uparrow$
Tensor Product	$\uparrow$	$\uparrow$	$\downarrow$
Zigzag Product	$\uparrow$	$\Downarrow$	$\downarrow$

Replacement Product

The replacement product $G \circledR H$ is the $2d$-degree graph defined by $\frac{1}{2}\widehat{A} + \frac{1}{2}(I_n \otimes B)$.

- G is an n-vertex regular graph of degree D, and A is the random walk matrix of G. H is a D-vertex regular graph of degree d, and B is the random walk matrix of H.

---

picture

If $\widehat{G}(u, l) = (v, m)$, place d parallel edges from the l-th vertex of H_u to the m-th vertex of H_v.

Lemma. $\lambda_{G \circledR H} \le 1 - \frac{(1-\lambda_G)(1-\lambda_H)^2}{24}$ and $\gamma_{G \circledR H} \ge \frac{1}{24}\gamma_G \gamma_H^2$.

Let A, B be the random walk matrices of G, H respectively.

$$\begin{aligned}
(A \circledR B)^3 &= \left(\frac{1}{2}\widehat{A} + \frac{1}{2}(I_n \otimes B)\right)^3 = \left(\frac{1}{2}\widehat{A} + \frac{1}{2}\left(I_n \otimes (\lambda_H E + \gamma_H J_D)\right)\right)^3 \\
&= \frac{1}{8}\left(\widehat{A} + \lambda_H (I_n \otimes E) + \gamma_H (I_n \otimes J_D)\right)^3 = \frac{1}{8}\left(\widehat{A}^3 + \ldots + \gamma_H^2 (I_n \otimes J_D)\widehat{A}(I_n \otimes J_D)\right) \\
&= \frac{1}{8}\left(\widehat{A}^3 + \ldots + \gamma_H^2 (A \otimes J_D)\right),
\end{aligned}$$

where the last equality is due to Lemma. Applying Lemma and the Claims, we get

$$(\lambda_{A \circledR B})^3 = \lambda_{(A \circledR B)^3} \le 1 - \frac{\gamma_H^2}{8} + \frac{\gamma_H^2}{8}\lambda_{A \otimes J_D} \le 1 - \frac{\gamma_H^2}{8} + \frac{\gamma_H^2}{8}\lambda_G = 1 - \frac{\gamma_H^2}{8}\gamma_G.$$

We have proved that $(\lambda_{G \circledR H})^3 \le 1 - \frac{\gamma_G \gamma_H^2}{8} \le \left(1 - \frac{\gamma_G \gamma_H^2}{24}\right)^3$. Hence $\gamma_{G \circledR H} \ge \frac{1}{24}\gamma_G \gamma_H^2$.

Expander Construction II

Theorem. There exists a strongly explicit $(4, \lambda)$-expander family for some $\lambda < 1$.

As a first step we prove that we can efficiently construct a family $\{G_k\}_k$ of graphs where each G_k has $(2d)^{100k}$ vertices.

1. Let H be a $((2d)^{100}, d, 0.01)$-expander graph, G_1 a $((2d)^{100}, 2d, 0.5)$-expander graph, and G_2 a $((2d)^{100\times 2}, 2d, 0.5)$-expander graph, all found by brutal force.
2. For $k > 2$ define
$$G_k = \left(G_{\lfloor \frac{k-1}{2} \rfloor} \otimes G_{\lceil \frac{k-1}{2} \rceil}\right)^{50} \textcircled{R} H.$$

G_k is a $((2d)^{100k}, 2d, 0.98)$-expander graph. [$k = 2^{\log k}$, hence the double exponential.]

Expander Construction II

Fact. G_k is a $((2d)^{100k}, 2d, 0.98)$-expander graph.

---

1. Let n_k be the number of vertices of G_k.

$$n_k = n_{\lfloor \frac{k-1}{2} \rfloor} n_{\lceil \frac{k-1}{2} \rceil} (2d)^{100} = (2d)^{100\lfloor \frac{k-1}{2} \rfloor} (2d)^{100\lceil \frac{k-1}{2} \rceil} (2d)^{100} = (2d)^{100k}.$$

2. $G_{\lfloor k-1 \rfloor}, G_{\lceil k-1 \rceil}$ degree $2d \Rightarrow G_{\lfloor k-1 \rfloor} \otimes G_{\lceil k-1 \rceil}$ degree $(2d)^2 \Rightarrow$ $(G_{\lfloor k-1 \rfloor} \otimes G_{\lceil k-1 \rceil})^{50}$ degree $(2d)^{100} \Rightarrow G_k$ degree $2d$.

3. $\lambda_{G_{\lfloor k-1 \rfloor}}, \lambda_{G_{\lceil k-1 \rceil}} \leq 0.98 \Rightarrow \lambda_{G_{\lfloor k-1 \rfloor} \otimes G_{\lceil k-1 \rceil}} \leq 0.98 \Rightarrow \lambda_{(G_{\lfloor k-1 \rfloor} \otimes G_{\lceil k-1 \rceil})^{50}} \leq 0.5 \Rightarrow$ $\lambda_{G_k} \leq 1 - 0.5(0.99)^2/24 < 0.98$.

Expander Construction II

There is a $\mathrm{poly}(k)$-time algorithm that upon receiving a label i of a vertex in G_k and an index j in $[2d]$ finds the j-th neighbor of i.

$$\begin{aligned}
\mathrm{time}(G_{k+1}) &= 50 \cdot \mathrm{time}(G_{\lfloor \frac{k-1}{2} \rfloor}) + 50 \cdot \mathrm{time}(G_{\lceil \frac{k-1}{2} \rceil}) + \mathrm{time}(H) \\
&= 2^{\log k} \cdot \mathrm{time}(G_2) + (2^{\log k - 1} + \ldots + 2 + 1) \cdot \mathrm{time}(H) \\
&= 2^{O(\log k)} \\
&= \mathrm{poly}(k).
\end{aligned}$$

The expander graph is strongly explicit.

Expander Construction II

Suppose $(2d)^{100k} < i < (2d)^{100(k+1)}$. Let $(2d)^{100(k+1)} = xi + r$.

- Divide the $(2d)^{100(k+1)}$ vertices into i classes among which r classes being of size $x+1$ and $i-r$ classes being of size x.
- Contract every class into a mega-vertex.
- Add $2d$ self-loops to each of the $i-r$ mega-vertices.

This is a $(i, 2d(x+1), (2d)0.01/(x+1))$ edge expander.

We get a $((2d)^{101}, 0.01/(2d)^{99})$ edge expander family.

Reingold's Theorem

Theorem. $\mathtt{UPATH} \in \mathbf{L}$.

---

1. O. Reingold. Undirected ST-Connectivity in Log-Space. STOC 2005.

The Idea

Connectivity Algorithm for d-degree expander graph is easy.

- The diameter of an expander graph is of length $O(\log(n))$.
- An exhaustive search can be carried out in logspace.

Reingold's idea is to transform conceptually a graph G to a graph G' so that a connected component in G becomes an expander in G' and unconnected vertices in G remain unconnected in G'.

Moreover finding a neighbor of a given vertex in the conceptual G' can be done in $O(\log|G|)$ space.

The Algorithm

Fix a $(d^{50}, d/2, 0.01)$-expander graph H for $d = 4$.

1. Convert the input graph G to a d^{50}-degree graph on the fly.
   1.1 Add self-loops to increase degree.
   1.2 Replace a large degree vertex by a cycle to decrease degree.
2. $G_0 = G$; $G_k = (G_{k-1} \circledR H)^{50}$ is constructed on the fly.
3. Apply Connectivity Algorithm to the expander $G_{10 \log n}$.

- If G_{k-1} is an N-vertex d^{50}-degree, $G_{k-1} \circledR H$ is an $d^{50}N$-vertex d-degree, and $(G_{k-1} \circledR H)^{50}$ is an $d^{50}N$-vertex d^{50}-degree.
- Conclude that $G_{10 \log n}$ contains $(d^{50})^{10 \log n} n = n^{1001}$ vertices.

The Complexity

Only paths of length $2\log_{\frac{1}{0.95}} n^{1001}$ need be considered.

1. Each vertex, except s, t, is coded up by $50\log(d)$ bits, say x, declaring that it is the 2^x-th neighbor of the previous vertex.
2. The algorithm keeps the current vertex. When backtracking it starts from s all over again to get the previous vertex.

Step 2 and Step 3 of the algorithm can be carried out on the fly using this mechanism.

- We cannot record all the vertices in a path. $[\log(n)^2]$

Lewis and Papadimitriou introduced **SL** as the class of problems solvable in logspace by an NTM that satisfies the following.

1. If the answer is 'yes,' one or more computation paths accept.
2. If the answer is 'no,' all paths reject.
3. If the machine can make a transition from configuration C to configuration D, then it can also goes from D to C.

Theorem. UPATH is **SL**-complete.

Corollary. UPATH is **L**-complete.

Proof.

Reingold Theorem implies that $\mathbf{L} = \mathbf{SL}$. □