Spectral Theory for Complex Networks

Charles Bordenave

CNRS & University of Toulouse

Overview

Consider a large network with $n \gg 1$ vertices.

Part I

There are natural matrices associated to the network.

Part II

(i) *Some geometric properties of the network can be recovered from the spectrum of these matrices.*

(ii) *Some algorithms or processes running on the network can be studied thanks to the spectrum of these matrices.*

Part III

In the regime $n \gg 1$, it is often possible to study the asymptotic properties of the spectrum.

Part I : Network Matrices

Matrices on the conductance model

Conductance Model

Let $V = \{1, \ldots, n\}$ and consider a collection of non-negative conductances

$$\{ X_{uv} : \; u, v \in V \}$$

X_{uv} *may for example represents an affinity of* u *for* v.

The associated directed graph $G = (V, E)$ is defined by $uv \in E$ if $X_{uv} > 0$.

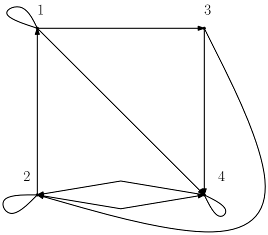
Adjacency and degree matrices

Define the $n \times n$ matrices

$$X = (X_{ij})_{1 \leqslant i,j \leqslant n} \quad \text{and}$$

$$D = \mathrm{diag}\left(\sum_{\ell} X_{1\ell}, \ldots, \sum_{\ell} X_{n\ell}\right).$$

$$X = \begin{pmatrix} 3 & 0 & 1 & 3 \\ 2 & 1 & 0 & 1 \\ 0 & 2 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \qquad D = \begin{pmatrix} 7 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$



Markov transition matrix

Associate a Markov chain on V with transition matrix P

$$P_{ij} = \frac{X_{ij}}{\sum_{\ell=1}^{n} X_{i\ell}}.$$

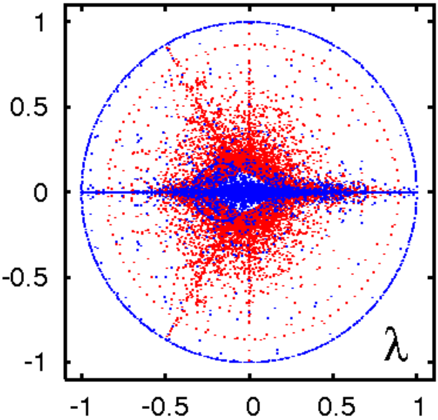
We have

$$P = D^{-1}X.$$

If $X_{uv} \in \{0,1\}$, P is the transition matrix of the random walk on G.

Google matrix : for $\alpha \in (0,1]$, $\alpha P + (1-\alpha)\mathbf{1}\mathbf{1}^*$.

Eigenvalues of transition matrix P for hypertext links of Cambridge University in 2006.



Frahm, Georgeot, Shepelyansky (2011).

Laplacian matrix

For $i \neq j$,

$$L_{ij} = X_{ij} \quad \text{and} \quad L_{ii} = -\sum_{\ell \neq i} X_{i\ell}.$$

The Laplacian L is the infinitesimal generator of the Markov process where the jump intensity from i to j is X_{ij}.

We have

$$L = X - D.$$

Normalized Laplacian : $D^{-1/2} L D^{-1/2} = D^{1/2}(P - I) D^{-1/2}$.

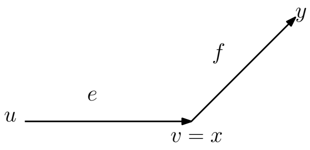
Non-Backtracking matrix

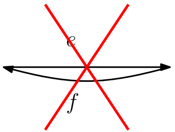
There are many other related matrices, more or less well understood.

If $e = uv, f = xy$ are in E,

$$B_{ef} = \mathbf{1}(v = x)\mathbf{1}(u \neq y),$$

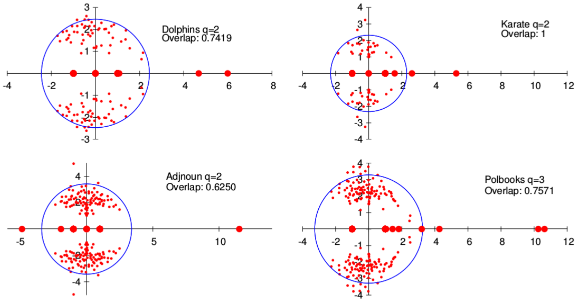
defines a $|E| \times |E|$ matrix on the oriented edges.





Non-Backtracking matrix

Used notably for community detection.



Krzakala/Moore/Mossel/Neeman/Sly/Zdeborová/Zhang (2013)

Part I : Network Matrices

Basic Properties of the Spectrum

Spectrum

For $A \in M_n(\mathbb{R})$, we denote its eigenvalues by

$$|\lambda_1(A)| \geqslant \ldots \geqslant |\lambda_n(A)|.$$

If A is symmetric, then the eigenvalues are real and there is an orthonormal basis of eigenvectors.

If $X_{uv} = X_{vu}$ then X and L are symmetric matrices and

$$P = D^{-1/2}(D^{-1/2} X D^{-1/2}) D^{1/2}$$

has also real eigenvalues.

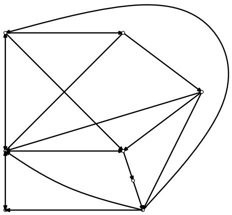
Perron-Frobenius Theorem

Assume that the graph G of X is strongly connected. Then X and P are said to be irreducible.

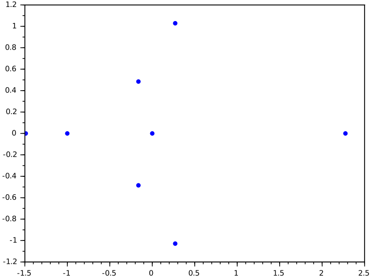
Then λ_1 is positive and is a simple eigenvalue. Its left and right eigenvector have positive coordinates.

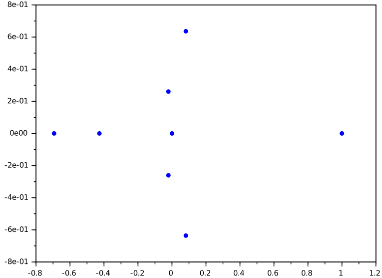
For the transition matrix P, $\lambda_1(P) = 1$ and $\pi P = \pi$ with $\sum_v \pi(v) = 1$ is the invariant probability measure of the Markov chain.

Example



$\pi \simeq (0.18, 0.06, 0.28, 0.22, 0.05, 0.03, 0.08, 0.11)$.





Laplacian matrix

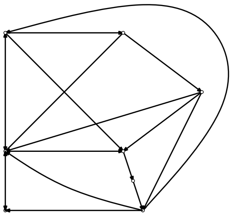
Assume that the graph G of X is strongly connected.

Then 0 is a simple eigenvalue of L. Its left and right eigenvector have positive coordinates.

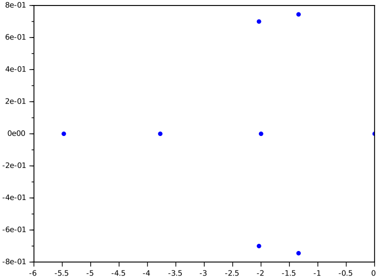
Then $\Pi L = 0$ with $\sum_v \Pi(v) = 1$ is the invariant probability measure of the Markov process. We have $\Pi \propto \pi D^{-1}$.

All other eigenvalues have negative real part (from Gershgorin Theorem, all eigenvalues are in $\cup_i B(-D_{ii}, D_{ii})$)

Example



$\Pi \simeq (0.18, 0.06, 0.28, 0.22, 0.05, 0.03, 0.08, 0.11)$.



Reversible case

When $X_{uv} = X_{vu}$, the graph G is undirected.

Then P and L are reversible processes and we find

$$\pi = \frac{1}{S}\left(D_{11}, \cdots, D_{nn}\right) \quad \text{and} \quad \Pi = \left(\frac{1}{n}, \cdots, \frac{1}{n}\right),$$

with

$$S = \sum_{u=1}^{n} D_{uu} = \sum_{u,v} X_{uv}.$$

Incidence matrices

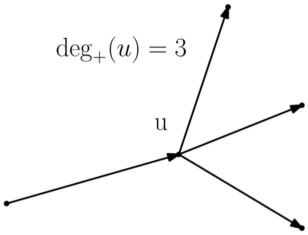
For simplicity, from now on

$$X_{uv} \in \{0,1\} \quad \text{and} \quad X_{uu} = 0.$$

Then

$$D_{uu} = \sum_{(uv)\in E} 1 = \deg_+(u),$$

is the outer degree of u.



Regular graphs

If for some d and for any $v \in V$,

$$\deg_+(v) = d.$$

then G is an outer-regular graph.

We have $D = dI$ and the matrices X, $L = dI - X$, $P = d^{-1}X$ have the same spectrum up to translation/dilation.

Also $\lambda_1(X) = d$ and the vector $\mathbf{1}$ is its eigenvector.

Part II : Spectrum and Graph Properties

Typical vs Extremal Eigenvalues

There are essentially two types of information encoded on the spectrum.

- the largest eigenvalues (and their eigenspaces) give some information on global graph properties (expansion, chromatic number, maximal cut, etc...),

- the typical eigenvalues give information on local graph properties (typical degree, partition function of spanning trees, matchings, etc...).

Part II : Spectrum and Graph Properties

Typical Eigenvalues and Partition Functions

Empirical spectral distribution (ESD)

The empirical distribution of the eigenvalues / density of states is the probability measure on $\mathbb{C}$,

$$\mu_A = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i(A)},$$

i.e. for any set $I \subset \mathbb{C}$

$$\mu_A(I) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(\lambda_i(A) \in I)$$

is the proportion of eigenvalues in I or equivalently, the probability that a typical eigenvalue is in I.

$$\int f d\mu_A = \frac{1}{n}\sum_{i=1}^{n} f(\lambda_i(A)).$$

Kirchoff Matrix-Tree Theorem

If G is an undirected connected graph then the number of spanning trees of G is equal to

$$t(G) = \frac{1}{n} \prod_{\lambda_i \neq 0} |\lambda_i|,$$

where $\lambda_i = \lambda_i(L)$.

In particular,

$$\frac{1}{n} \log t(G) = \int_{-\infty}^{0^-} \log |\lambda| d\mu_L(\lambda) - \frac{1}{n} \log n.$$

Closed paths

For t integer, let

$$S_t \quad = \quad |\{\text{closed paths of length } t \text{ in } G\}|$$

We have

$$S_t = \mathrm{Tr}\{X^t\} = \sum_{i=1}^{n} \lambda_i(X)^t = n \int \lambda^t d\mu_X.$$

In particular, for $z \in \mathbb{C}$, $\mathfrak{Im}(z) > 0$,

$$\frac{1}{n} \sum_{t \geqslant 0} \frac{S_t}{z^{t+1}} = \sum_{t \geqslant 0} \int \frac{\lambda^t}{z^{t+1}} d\mu_X = \int \frac{1}{z - \lambda} d\mu_X$$

is the Cauchy-Stieltjes transform of μ_X.

Return times

If Z_t is the Markov chain with transition matrix P,

$$\frac{1}{n}\sum_{v=1}^{n}\mathbb{P}(Z_t = v | Z_0 = v) = \frac{1}{n}\mathrm{Tr}\{P^t\} = \int x^t d\mu_P.$$

Similarly, for $t > 0$ real, if Z_t is the Markov process with generator L,

$$\frac{1}{n}\sum_{v=1}^{n}\mathbb{P}(Z_t = v | Z_0 = v) = \int e^{tL} d\mu_L.$$

Part II : Spectrum and Graph Properties

Eigenvectors

Localization/Delocalization of eigenvectors

Take X or L in the reversible case. Let $\psi_1, \dots \psi_n$ be an orthonormal basis of eigenvectors.

Then, for any $k \geqslant 1$,

$$\left(|\psi_k(1)|^2, \dots, |\psi_k(n)|^2\right)$$

is a probability vector on V.

Eigenvectors are of prime importance for the study of quantum mechanics in disorder media *Anderson (1956)*, quantum percolation *De Gennes, Lafore, Millot (1957)*.

Eigenvectors of the largest eigenvalues of the Google matrix are also studied, e.g. *Frahm, Georgeot, Shepelyansky (2011)*.

Delocalization of eigenvectors

How far is a typical eigenvector from a uniform vector on $S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$?

For example, do we have quantum ergodicity, i.e.

$$\frac{1}{\sum_k \mathbf{1}(\lambda_k \in I)} \sum_k \mathbf{1}(\lambda_k \in I) \sum_{v=1}^{n} f(v) |\psi_k(v)|^2 \simeq \frac{1}{n} \sum_{v=1}^{n} f(v).$$

Part II : Spectrum and Graph Properties

Local Notion of Spectrum

Spectral measure at a vector

Let A be a symmetric matrix, (e.g. X or L in the reversible case). Let $\psi_1, \ldots \psi_n$ be an orthonormal basis of eigenvectors.

Take $1 \leqslant x \leqslant n$, define the probability measure,

$$\mu_A^x = \sum_{k=1}^{n} |\psi_k(x)|^2 \delta_{\lambda_k}.$$

We have

$$\int \lambda^t d\mu_A^x = (A^t)_{xx},$$

and

$$\frac{1}{n}\sum_{x=1}^{n} \mu_A^x = \frac{1}{n}\sum_{x=1}^{n}\sum_{k=1}^{n} |\psi_k(x)|^2 \delta_{\lambda_k} = \frac{1}{n}\sum_{k=1}^{n} \delta_{\lambda_k} \sum_{x=1}^{n} |\psi_k(x)|^2 = \mu_A.$$

Part II : Spectrum and Graph Properties

Extremal Eigenvalues and Expanders

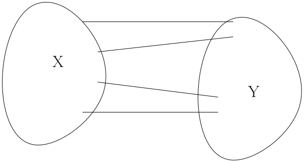
CHEEGER'S CONSTANT

Assume that G is undirected and connected.

For $X, Y \subset V$, define

$$\mathrm{vol}(X) = \sum_{x \in X} \deg(x).$$

$$E(X, Y) = \sum_{x \in X, y \in Y} \mathbf{1}(uv \in E).$$



Isoperimetric / Expansion constant :

$$h(G) = \min_{X \subset V} \frac{E(X, X^c)}{\min\left(\mathrm{vol}(X), \mathrm{vol}(X^c)\right)}.$$

Cheeger's Inequality

Let

$$1 = \lambda_1 > \lambda_2 \geqslant \cdots \geqslant \lambda_n \geqslant -1$$

be the eigenvalues of P.

$1 - \lambda_2$ is called the spectral gap of P.

Theorem

$$\frac{h(G)^2}{2} \leqslant 1 - \lambda_2 \leqslant 2h(G).$$

Cheeger's Inequality

Since

$$P = D^{-1}X = D^{-1/2}(D^{-1/2}XD^{-1/2})D^{1/2},$$

the λ_i 's are also the eigenvalues of S with $S = D^{-1/2}XD^{-1/2}$.

Since $P\mathbf{1} = \mathbf{1}$, $\chi = D^{1/2}\mathbf{1}$ is the eigenvector of S associated to $\lambda_1 = 1$.

Hence, from Courant-Fisher formula,

$$\lambda_2 = \max_{x:\langle x,\chi\rangle=0} \frac{\langle Sx, x\rangle}{\|x\|^2}.$$

Or equivalently,

$$1 - \lambda_2 = \min_{x:\langle x,\chi\rangle=0} \frac{\langle (I-S)x, x\rangle}{\|x\|^2}.$$

We set $\pi(x) = \deg(x) = (D\mathbf{1})(x)$. Since $\pi = D^{1/2}\chi$ with the change of variable $f = D^{-1/2}x$, after some algebra,

$$1 - \lambda_2 = \min_{f:\langle f,\pi\rangle=0} \frac{\sum_{u\sim v}(f(u)-f(v))^2}{\sum_v \deg(v) f(v)^2}.$$

Let X such that

$$h(G) = \frac{E(X, X^c)}{\min\left(\mathrm{vol}(X), \mathrm{vol}(X^c)\right)}.$$

We take

$$f(v) = \frac{\mathbf{1}(v \in X)}{\mathrm{vol}(X)} - \frac{\mathbf{1}(v \notin X)}{\mathrm{vol}(X^c)}.$$

We have

$$\langle f, \pi \rangle = \sum_{x \in X} \frac{\deg(x)}{\mathrm{vol}(X)} - \sum_{x \in X^c} \frac{\deg(x)}{\mathrm{vol}(X^c)} = 0,$$

and

$$\begin{aligned}
1 - \lambda_2 &\leqslant \frac{\sum_{u \sim v} (f(u) - f(v))^2}{\sum_v \deg(v) f(v)^2} \\
&= 2E(X, X^c) \frac{(1/\mathrm{vol}(X) - 1/\mathrm{vol}(X^c))^2}{1/\mathrm{vol}(X) + 1/\mathrm{vol}(X^c)} \\
&\leqslant 2 \frac{E(X, X^c)}{\min(\mathrm{vol}(X), \mathrm{vol}(X^c))} \\
&\leqslant 2h(G).
\end{aligned}$$

Alon-Boppana bound

If G is a d-regular undirected graph on n vertices, then $\lambda_1(X) = d$ and

$$\lambda_2(X) \geqslant 2\sqrt{d-1} + o(1).$$

Consequently,

$$1 - \lambda_2(P) \leqslant 1 - 2\frac{\sqrt{d-1}}{d} + o(1).$$

Ramanujan graphs

Let G be a d-regular undirected graph.

G is Ramanujan if

$$\lambda_2(X) \leqslant 2\sqrt{d-1}.$$

They are the best possible expanders. There is a generalized definition for non-regular graphs (lifts of graphs).

Existence of Ramanujan graphs

Sequence of Ramanujan graphs $G_1, G_2, \cdots$, with $|V(G_n)|$ growing to infinity, are known to exist when

- $d = q + 1$ with $q = p^k$ and p prime number *Lubotzky, Phillips & Sarnak (1988), Morgenstern (1994).*

- any $d \geqslant 3$, *Marcus, Spielman, Srivastava (2013).*

Part II : Spectrum and Graph Properties

Extremal Eigenvalues and Combinatorial Optimization

Assume that G is undirected and connected

If Δ is the maximal degree and $\overline{\deg}$ the average degree

$$\max(\sqrt{\Delta}, \overline{\deg}) \leqslant \lambda_1(X) \leqslant \Delta.$$

If χ is the chromatic number and $\lambda_n(X) \leqslant \ldots \leqslant \lambda_1(X)$,

$$1 - \frac{\lambda_1(X)}{\lambda_n(X)} \leqslant \chi \leqslant 1 + \lambda_1(X).$$

...

Part II : Spectrum and Graph Properties

Spectral Gaps and Convergence to Equilibrium

Spectral gap

The spectrum of P and L contain many information on the behavior of the Markov chain/process.

Notably through the spectral gap

$$-\max_{\lambda \neq 0} \mathfrak{Re} \lambda(L)$$

$$1 - \max_{\lambda \neq 1} \mathfrak{Re} \lambda(P)$$

Even more connections for reversible Markov chain/process. For simplicity we only consider L.

Spectral gap

Assume that X is reversible. Let Z_t be a Markov process with generator L,

$$P_t^x = e^{tL} e_x$$

is the probability distribution of Z_t given $Z_0 = x$.

Let $\lambda_1 = 0 > \lambda_2 \geqslant \cdots \geqslant \lambda_n$ the eigenvalues of L and $\psi_1 = \mathbf{1}/\sqrt{n}, \ldots, \psi_n$ an orthogonal basis of eigenvectors.

From the spectral theorem

$$\begin{aligned} e^{tL} &= \sum_{i=1}^{n} e^{t\lambda_i} \psi_i \psi_i^* \\ P_t^x &= \frac{1}{n} + \sum_{i=2}^{n} e^{t\lambda_i} \psi_i \psi_i(x) \end{aligned}$$

Spectral gap

Recall that $\Pi = \mathbf{1}/n$ is the invariant distribution. We get

$$\|P_t^x - \Pi\|^2 = \sum_{i=2}^{n} e^{2t\lambda_i} |\psi_i(x)|^2 \leqslant e^{-2|\lambda_2|t}$$

Recall

$$\|x\| \leqslant \sum_i |x_i| \leqslant \sqrt{n}\|x\|.$$

So,

$$|\psi_2(x)| e^{-|\lambda_2|t} \leqslant 2\|P_t^x - \Pi\|_{TV} \leqslant \sqrt{n} e^{-|\lambda_2|t}.$$

where the total variation norm is

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_x |\mu(x) - \nu(x)|.$$

Spectral gap

The mixing time of a Markov process is usually defined as

$$\tau = \inf_{t>0} \max_x \|P_t^x - \Pi\|_{TV} \leqslant \frac{1}{2}.$$

$$\frac{\max_x |\psi_2(x)|}{|\lambda_2|} \leqslant \tau \leqslant \frac{\log n}{2|\lambda_2|}.$$

(Note that $\max_x |\psi_2(x)| \geqslant 1/\sqrt{n}$).

There are similar developments for reversible Markov chains and also partially in the non-reversible case.

(Levin/Peres/Wilmer 2009).

Part III : Spectrum of large random graphs

Simple models of random graphs

Average Degree

The number of directed edges is

$$|E| = \sum_{v \in V} \deg_+(v) = \sum_{v \in V} \deg_-(v) = -\mathrm{Tr}\{L\}.$$

Hence

$$\overline{\deg}(G) = \frac{|E|}{n}$$

is the average degree of a vertex.

Three Rough Classes of Graphs

Let $G = (V, E)$ with $n = |V| \gg 1$. We say that G is

Dense if

$$\overline{\text{deg}}(G) = \Theta(n).$$

Sparse if

$$1 \ll \overline{\text{deg}}(G) = o(n).$$

Diluted if

$$\overline{\text{deg}}(G) = O(1).$$

Three Rough Classes of Graphs

The spectrum of large dense and sparse random graphs can be studied with tools from random matrix theory.

The spectrum of large diluted random graphs can be studied (hopefully) with tools from random Schrödinger operators.

Erdős-Rényi random graph

Taking $p \in [0,1]$. A random variable X has a $\mathrm{Ber}(p)$ law if

$$\mathbb{P}(X = 1) = p = 1 - \mathbb{P}(X = 0).$$

- Reversible model : $(X_{ij})_{1 \leqslant i < j \leqslant n}$ are independent $\mathrm{Ber}(p)$ variables and $X_{ij} = X_{ji}$.

- Non-reversible model : $(X_{ij})_{1 \leqslant i \neq j \leqslant n}$ are independent $\mathrm{Ber}(p)$ variables.

More generally, we could consider a general dependence of (X_{ij}, X_{ji}).

We have

$$\mathbb{E}\deg(v) = \mathbb{E}\sum_{u \neq v} X_{uv} = (n-1)p.$$

$$p \in (0,1) : \text{dense}$$

$$np \to \infty,\ p = o(1) : \text{sparse}$$

$$p = c/n : \text{diluted.}$$

Inhomogeneous random graphs

Let $W : [0,1]^2 \to [0,1]$ be a constant by block function (independent of n). For example,

$$W([0,1]^2) = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}.$$

For the non-reversible case, we set $(X_{ij})_{1 \leqslant i,j \leqslant n}$ are independent $\mathrm{Ber}(p_{ij})$ variables with

$$p_{ij} = p\, W\left(\frac{i}{n}, \frac{j}{n}\right),$$

and $p = p(n)$ gives the order of the average degree.

In a sense, for $p = 1$, these graphs are dense among dense graphs.

Uniform d-regular graphs

Take dn even, we may define G as a random d-regular sampled uniformly at random (provided the set is not-empty).

If $d = o(\sqrt{n})$, G can be studied thanks to a simple probabilistic model, the configuration model, *Bollobás (1981)*.

The configuration model can also be used to study uniform random graphs with given degree sequences.

Why studying toy models ?

Very simple models of random networks : no underlying geometry.

Their study may help to understand what is due to noise in a real world network.

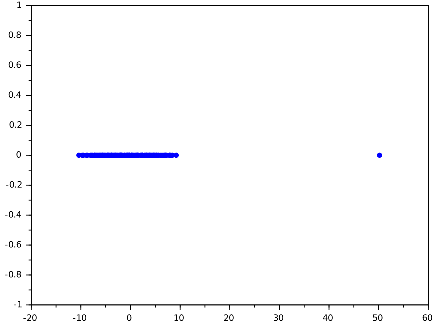
They are natural candidates for best expanders, best mixing times, graphs with delocalized eigenvectors, etc.

Part III : Spectrum of large random graphs

Dense undirected random graphs

Dense Erdős-Rényi

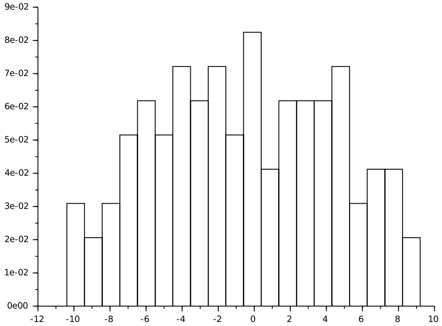
We fix $p \in (0,1)$.
Reversible model : $(X_{ij})_{1 \leqslant i < j \leqslant n}$ are independent random variables $\mathrm{Ber}(p)$ and $X_{ij} = X_{ji}$.



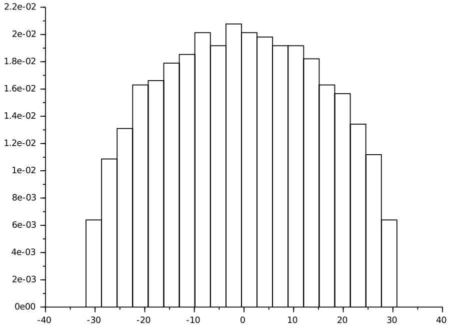
Eigenvalues of X for $n = 100$, $p = 1/2$.

Dense Erdős-Rényi

Histogram of eigenvalues $\lambda_2(X) \geqslant \ldots \geqslant \lambda_n(X)$ for $n = 100$ and $p = 1/2$.



Histogram of eigenvalues $\lambda_2(X) \geqslant \ldots \geqslant \lambda_n(X)$ for $n = 1000$ and $p = 1/2$.



Wigner's semi-circle law

Wigner matrix :

$$Y = (Y_{ij})_{1 \leqslant i,j \leqslant n}$$

with $(Y_{ij})_{i>j}$ iid random variables, independent of $(Y_{ii})_{i \geqslant 1}$ iid and $Y_{ji} = Y_{ij}$,

Theorem
If $\mathbb{E}Y_{11} = \mathbb{E}Y_{12} = 0$, $\mathbb{E}Y_{12}^2 = 1$ *then with probability one, for any interval* $I \subset \mathbb{R}$,

$$\mu_{\frac{Y}{\sqrt{n}}}(I) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left(\frac{\lambda_i(Y)}{\sqrt{n}} \in I \right) \to \mu_{sc}(I).$$

where $\mu_{sc}(dx) = \mathbf{1}_{|x| \leqslant 2} \sqrt{4 - x^2}dx$. *In other words, a.s.* $\mu_{Y/\sqrt{n}}$ *converges weakly to* μ_{sc}.

Second moment

Assume $Y_{ii} = 0$.

$$\begin{aligned}\int x^2 d\mu_{\frac{Y}{\sqrt{n}}} &= \frac{1}{n}\mathrm{Tr}\left(\frac{Y}{\sqrt{n}}\right)^2 \\ &= \frac{1}{n^2}\sum_{i,j} Y_{ij}^2 \\ &= \frac{2}{n^2}\sum_{1\leqslant i<j\leqslant n} Y_{ij}^2\end{aligned}$$

Taking expectation, from $\mathbb{E}Y_{12}^2 = 1$,

$$\mathbb{E}\int x^2 d\mu_{\frac{Y}{\sqrt{n}}} = \frac{n(n-1)}{n^2} = 1 + o(1).$$

Hence, from the law of large numbers, with probability one,

$$\int x^2 d\mu_{\frac{Y}{\sqrt{n}}} \to 1 + o(1).$$

Method of moments

We have

$$\int x^{2k+1} d\mu_{sc} = 0 \quad \text{and} \quad \int x^{2k} d\mu_{sc} = c_k,$$

where c_k is the k-th Catalan number

$$c_k = \frac{1}{k+1}\binom{2k}{k}.$$

By a combinatorial argument, we can prove that

$$\int x^k d\mu_{\frac{Y}{\sqrt{n}}} = \frac{1}{n}\mathrm{Tr}\left(\frac{Y}{\sqrt{n}}\right)^k$$

$$= \frac{1}{n^{1+k/2}} \sum_{i_1,\cdots,i_k} \prod_{\ell=1}^{k} Y_{i_\ell i_{\ell+1}}$$

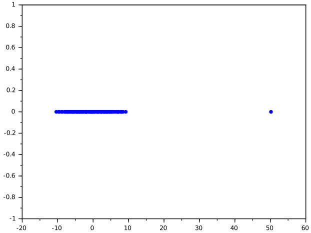
$$= \int x^k d\mu_{sc} + o(1).$$

Semi-circle law for adjacency matrix

Theorem

Fix $p \in (0,1)$ *and let* $\sigma^2 = p(1-p)$. *With probability one, for any interval* $I \subset \mathbb{R}$,

$$\mu_{\frac{X}{\sigma\sqrt{n}}}(I) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\left(\frac{\lambda_i(X)}{\sigma\sqrt{n}} \in I\right) \to \mu_{sc}(I).$$



From Wigner matrix to Adjacency matrix

We have

$$\sigma^2 = \mathrm{Var}(X_{12}) = \mathbb{E}(X_{12} - \mathbb{E}X_{12})^2 = \mathbb{E}(X_{12}^2) - (\mathbb{E}X_{12})^2 = p(1-p).$$

Also, with $J = \mathbf{1}\mathbf{1}^*$,

$$\mathbb{E}X = p(J - I).$$

Now,

$$\frac{X}{\sigma\sqrt{n}} = \frac{(X - \mathbb{E}X)}{\sigma\sqrt{n}} + \frac{pJ}{\sigma\sqrt{n}} - \frac{pI}{\sigma\sqrt{n}}.$$

From Wigner matrix to Adjacency matrix

$$\frac{X}{\sigma\sqrt{n}} = \frac{(X - \mathbb{E}X)}{\sigma\sqrt{n}} + \frac{pJ}{\sigma\sqrt{n}} - \frac{pI}{\sigma\sqrt{n}}.$$

- $\frac{(X-\mathbb{E}X)}{\sigma\sqrt{n}}$ is a Wigner matrix, its ESD converge to the semi-circular law.

- $\frac{pJ}{\sigma\sqrt{n}}$ has rank one (but norm of order $\sqrt{n}$!).

- $\frac{pI}{\sigma\sqrt{n}}$ has norm of order $1/\sqrt{n}$ (but full rank !).

Interlacing Inequalities

Take $A \in M_n(\mathbb{R})$ symmetric and $A' \in M_{n-1}(\mathbb{R})$ a minor,

$$A = \begin{pmatrix} A' & v \\ v^* & u \end{pmatrix}$$

Then, if $\lambda_{i+1} \leqslant \lambda_i$,

$$\lambda_{i+1}(A) \leqslant \lambda_i(A') \leqslant \lambda_i(A).$$

Take $A, B \in M_n(\mathbb{R})$ symmetric and $\mathrm{rank}(A - B) = k$ then

$$\lambda_{i+k}(A) \leqslant \lambda_i(B) \leqslant \lambda_{i-k}(A).$$

Interlacing Inequalities

Define, the Kolmogorov-Smirnov distance as

$$d_{KS}(\mu, \nu) = \sup_t |\mu((-\infty, t)) - \nu((-\infty, t))|.$$

If $d_{KS}(\mu_n, \mu) \to 0$ then μ_n converges weakly to μ.

Lemma
If A, B are Hermitian $n \times n$ matrices, then

$$d_{KS}(\mu_A, \mu_B) \leqslant \frac{\operatorname{rank}(A-B)}{n}.$$

From Wigner matrix to Adjacency matrix

$$\frac{X}{\sigma\sqrt{n}} = \frac{(X - \mathbb{E}X)}{\sigma\sqrt{n}} + \frac{pJ}{\sigma\sqrt{n}} - \frac{pI}{\sigma\sqrt{n}}.$$

- $\frac{(X-\mathbb{E}X)}{\sigma\sqrt{n}}$ is a Wigner matrix, its ESD converge to the semi-circular law.

- $\frac{pJ}{\sigma\sqrt{n}}$ has rank one (but norm of order $\sqrt{n}$!).

- $\frac{pI}{\sigma\sqrt{n}}$ has norm of order $1/\sqrt{n}$ (but full rank!).

A similar study can be done for the ESD of dense undirected inhomogeneous graphs, where

$$X_{ij} \stackrel{d}{=} \mathrm{Ber}(W(i/n, j/n)),$$

and $W : [0,1]^2 \to [0,1]$ is a constant by block function.

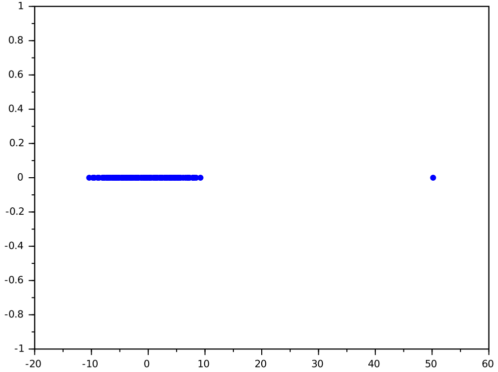
The limit is the semi-circle if

$$\sigma^2(x) = \int_0^1 W(x,y)(1 - W(x,y))dy$$

is a constant funciton of x.

Extremal eigenvalues in dense Erdős-Rényi

We fix $p \in (0,1)$. $(X_{ij})_{1 \leqslant i < j \leqslant n}$ are independent random variables $\mathrm{Ber}(p)$ and $X_{ij} = X_{ji}$.



Eigenvalues of X for $n = 100$, $p = 1/2$.

Fűredi-Komlós Theorem

Wigner matrix

$$Y = (Y_{ij})_{1 \leqslant i,j \leqslant n}$$

with $(Y_{ij})_{i>j}$ iid random variables, independent of $(Y_{ii})_{i \geqslant 1}$ iid and $Y_{ji} = Y_{ij}$,

Theorem
If $\mathbb{E}Y_{11} = \mathbb{E}Y_{12} = 0$, $\mathbb{E}Y_{12}^2 = 1$, $\mathbb{E}Y_{11}^2 < \infty$, $\mathbb{E}Y_{11}^4 < \infty$, *then with probability one,*

$$\lambda_1\left(\frac{Y}{\sqrt{n}}\right) = 2 + o(1) = -\lambda_n\left(\frac{Y}{\sqrt{n}}\right).$$

Recall that $\text{supp}(\mu_{sc}) = [-2, 2]$.

Theorem

Fix $p \in (0,1)$ *and let* $\sigma^2 = p(1-p)$. *With probability one,*

$$\lambda_1(X) = pn + O(\sqrt{n}),$$

$$\lambda_2(X) = 2\sigma\sqrt{n} + o(\sqrt{n}) = -\lambda_n(X),$$

and, if ψ_1 *is the Perron eigenvector,*

$$\|\psi_1 - \mathbf{1}/\sqrt{n}\| = O(1/\sqrt{n}).$$

Bauer-Fike Theorem

Theorem
If A, B *is* $n \times n$ *Hermitian,*

$$|\lambda_i(A+B) - \lambda_i(A)| \leqslant \|B\|.$$

For general B*, for some permutation* σ*,*

$$\left|\lambda_i(A+B) - \lambda_{\sigma(i)}(A)\right| \leqslant \|B\|.$$

$$X = p(J - I) + (X - \mathbb{E}X).$$

The eigenvalues of $J - I$ are : $n - 1$ and -1 with multiplicity $n - 1$.

From Fűredi-Komlós Theorem, $\|X - \mathbb{E}X\| = 2\sigma\sqrt{n} + o(\sqrt{n})$

Hence,

$$\begin{aligned} |\lambda_1(X) - p(n-1)| &= O(\sqrt{n}) \\ |\lambda_2(X) + p| &\leqslant (2 + o(1))\sigma\sqrt{n}. \end{aligned}$$

However, from the semi-circle law, we already know that $\lambda_2(X) \geqslant (2 + o(1))\sigma\sqrt{n}$.

Rank one perturbation of Eigenvector

Assume that

$$B = A + uv^*.$$

If λ is an eigenvalue of B and not of A then

$$1 + v^*(A - \lambda)^{-1}u = 0$$

and

$$\psi = (A - \lambda)^{-1}u.$$

is an eigenvector.

Rank one perturbation of Eigenvector

Apply this

$$X = (X - \mathbb{E}X - pI) + p\mathbf{1}\mathbf{1}^*,$$

and

$$\lambda = \lambda_1(X) = pn + O(\sqrt{n}).$$

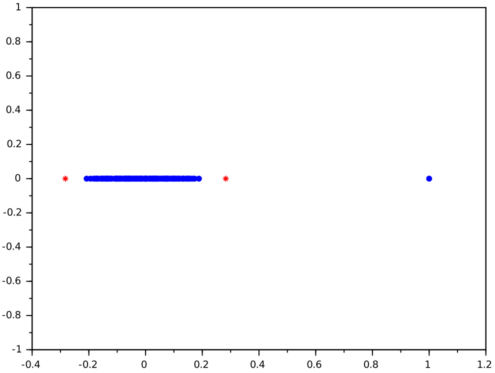
From $\|X - \mathbb{E}X - pI\| = O(\sqrt{n})$, we get

$$\|\psi_1 - \mathbf{1}/\sqrt{n}\| = O(1/\sqrt{n}).$$

Markov Transition Matrix

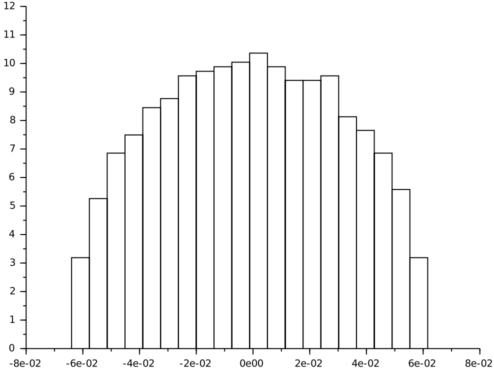
A similar analysis can be done for $P = D^{-1}X$.

$$\lambda_2(P) = \frac{2\sigma}{p\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right) = -\lambda_n(P).$$



Eigenvalues of P for $n = 100$, $p = 1/2$.

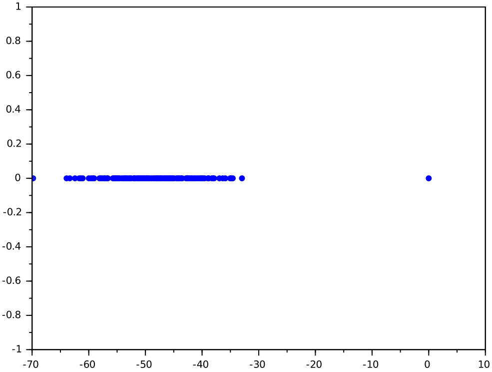
Markov Transition Matrix



Histogram of $\lambda_n(P) \leqslant \ldots \leqslant \lambda_2(P)$ for $n = 1000$, $p = 1/2$.

Laplacian

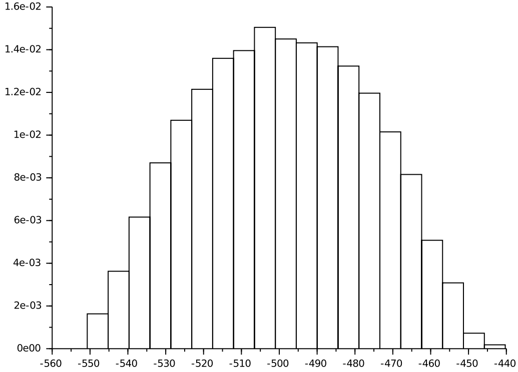
$$L = X - D.$$



Eigenvalues of L for $n = 100$, $p = 1/2$.

LAPLACIAN

$$L = X - D.$$



Histogram of $\lambda_n(X) \leqslant \ldots \leqslant \lambda_2(L)$ for $n = 1000$, $p = 1/2$.

Laplacian

Theorem

Fix $p \in (0,1)$ *and let* $\sigma^2 = p(1-p)$. *With probability one, for any interval* $I \subset \mathbb{R}$,

$$\mu_{\frac{L+npI}{\sigma\sqrt{n}}}(I) \to \mu(I).$$

where $\mu = \mu_{sc} \boxplus N(0,1)$. *Moreover, with probability one,*

$$\lambda_2(L) = -np + \sigma(1+o(1))\sqrt{2n\log n}.$$

(Ding/Jiang 2010), (Jiang 2012)

Heuristics

$$\frac{L+npI}{\sigma\sqrt{n}} \quad = \quad \frac{X-\mathbb{E}X}{\sigma\sqrt{n}} - \frac{D-\mathbb{E}D}{\sigma\sqrt{n}} + \frac{pJ}{\sigma\sqrt{n}}.$$

- $\frac{X-\mathbb{E}X}{\sigma\sqrt{n}}$: Wigner matrix : its ESD converges to the semi-circular law.

- $\frac{D-\mathbb{E}D}{\sigma\sqrt{n}}$: diagonal matrix with approximately iid Gaussian $N(0,1)$ coefficients, $D_{ii} = \sum_j X_{ij}$.

- $\frac{pJ}{\sigma\sqrt{n}}$: one eigenvalue $\frac{p\sqrt{n}}{\sigma} \to \infty$, all others 0.

Free convolution

Let A_n be a sequence of deterministic Hermitian $n \times n$ matrices such that for any bounded continuous function f,

$$\int f\mu_{A_n} \to \int f d\mu.$$

Then, if Y is a Wigner matrix, with probability one,

$$\int f\mu_{\frac{Y}{\sqrt{n}}+A_n} \to \int f d\nu.$$

and

$$\nu := \mu_{sc} \boxplus \mu.$$

In high dimension, the spectra add up!!

Maximum of Gaussian variables

$$\frac{L+npI}{\sigma\sqrt{n}} \quad = \quad \frac{X-\mathbb{E}X}{\sigma\sqrt{n}} - \frac{D-\mathbb{E}D}{\sigma\sqrt{n}} + \frac{pJ}{\sigma\sqrt{n}}.$$

If $(Z_i)_{i\geqslant 1}$ are iid $N(0,1)$ variables then with probability one,

$$\max_{1\leqslant i\leqslant n} Z_i = (1+o(1))\sqrt{2\log n}.$$

In particular,

$$\left\|\frac{D-\mathbb{E}D}{\sigma\sqrt{n}}\right\| = (1+o(1))\sqrt{2\log n},$$

and ...

$$\lambda_2(L) = -np + \sigma(1+o(1))\sqrt{2n\log n}.$$

Part III : Spectrum of large random graphs

Sparse undirected random graphs

Sparse Erdős-Rényi

The above results remain valid as long as

$$np \to \infty,$$

for the ESD of X or

$$\frac{np}{\log n} \to \infty.$$

for all other statements.

Note that $\sigma = \sqrt{p(1-p)} \sim \sqrt{p}$ when $p = o(1)$.

A key technical statement is *Khorunzhy (2001), Vu (2007)* : with probability one,

$$\left\| \frac{X - \mathbb{E}X}{\sigma\sqrt{n}} \right\| = 2 + o(1)$$

when

$$\frac{np}{\log n} \to \infty.$$

However, when $p = o\left(\frac{\log n}{n}\right)$, $\left\| \frac{X - \mathbb{E}X}{\sigma\sqrt{n}} \right\| \gg 1$.

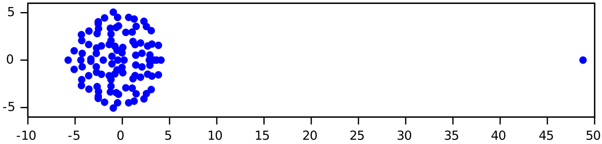
Part III : Spectrum of large random graphs

Sparse/dense directed random graphs

Directed Erdős-Rényi

Irreversible model : $(X_{ij})_{1\leqslant i\neq j\leqslant n}$ are independent random variables $\mathrm{Ber}(p)$.

Eigenvalues of X for $n = 100$, $p = 1/2$.



Circular Law

Theorem
Assume that $(Y_{ij})_{i,j\geqslant 1}$ *are iid random variables* $\mathbb{E}Y_{12}=0$, *and* $\mathbb{E}|Y_{12}|^2=1$, *consider the matrix*

$$Y=(Y_{ij})_{1\leqslant i,j\leqslant n}.$$

Then, with probability one, for any Borel set $I\subset\mathbb{C}$,

$$\mu_{Y/\sqrt{n}}(I)\to\mu_c(I),$$

with

$$\mu_c(dxdy)=\frac{1}{\pi}\mathbf{1}_{|z|\leqslant 1}dz.$$

(Mehta 1967), (Girko 1984), (Bai 1997), (Pan/Zhou 2010), (Götze/Tikhomirov 2010) ..., (Tao/Vu 2010).

Circular law for adjacency matrix

Theorem
Assume $(\log n)^6/n \ll p \leqslant 1-\delta$ *and let* $\sigma^2 = p(1-p)$. *For any Borel* $I \subset \mathbb{C}$, *in probability,*

$$\mu_{\frac{X}{\sigma\sqrt{n}}}(I) \to \mu_c(I).$$

(Bordenave, Caputo, Chafaï 2014)

Heuristics

$$\frac{X}{\sigma\sqrt{n}} = \frac{(X - \mathbb{E}X)}{\sigma\sqrt{n}} + \frac{pJ}{\sigma\sqrt{n}} - \frac{pI}{\sigma\sqrt{n}}.$$

- $\frac{(X-\mathbb{E}X)}{\sigma\sqrt{n}}$ is a random iid matrix, its ESD converges to the circular law.
- $\frac{pJ}{\sigma\sqrt{n}}$ has rank one (but norm of order $\sqrt{n}$!).
- $\frac{pI}{\sigma\sqrt{n}}$ has norm of order $1/\sqrt{n}$ (but full rank!).

Perturbation of non-Hermitian matrices

Take

$$N = \begin{pmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \ddots & \ddots \\ 0 & 0 & \dots & \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \ddots & \ddots \\ \varepsilon & 0 & \dots & \end{pmatrix}.$$

All eigenvalues of N are 0,

$$\mu_N = \delta_0.$$

For $\varepsilon = 1$, the eigenvalues of C are $e^{2i\pi k/n}$, $1 \leqslant k \leqslant n$:

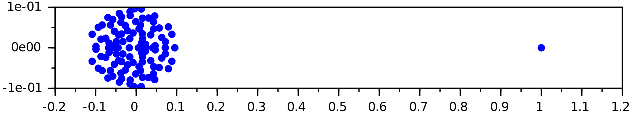
$$\int f d\mu_C \to \frac{1}{2\pi} \int_0^{2\pi} f(e^{2i\theta}) d\theta.$$

True as soon as $\varepsilon^{1/n} \to 1$!!

Markov Transition Matrix

For $(\log n)^6/n \ll p \leqslant 1-\delta$, there is a circular law for $P = D^{-1}X$ with radius $\sigma/(p\sqrt{n})$.

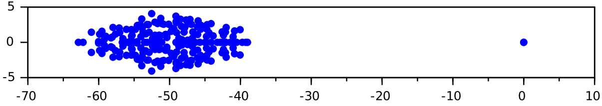
Eigenvalues of P for $n = 100$, $p = 1/2$.



Laplacian

In the same regime, there is a convergence for the ESD of $L = X - D$ shifted by pn and rescaled by $\sigma\sqrt{n}$.

Eigenvalues of L for $n = 100$, $p = 1/2$.



Theorem.
If $p \gg (\log n)/n$, *then, a.s. for* $n \gg 1$, *the Markovian generator* L *is irreducible and*

$$\|\Pi - \mathbf{1}/n\|_{\mathrm{TV}} = O\left(\frac{\sigma}{p}\sqrt{\frac{\log n}{n}}\right) + O\left(\frac{\sqrt{\sigma}}{p}\frac{\log n}{n^{3/4}}\right).$$

Similarly, for the Markov transition matrix P,

$$\|\pi - \mathbf{1}/n\|_{\mathrm{TV}} = O\left(\frac{\sigma}{p\sqrt{n}}\right) + O\left(\frac{\sqrt{\sigma \log n}}{pn^{3/4}}\right),$$

(Bordenave, Caputo, Chafaï 2014)

Part III : Spectrum of large random graphs

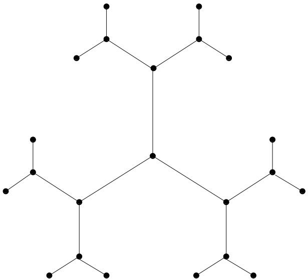
Diluted undirected random graphs

Kesten-McKay Law

Fix integer $d \geqslant 1$. Let $G = G_n$ be a sequence of d-regular graphs on n vertices such that for any k,

$$|\{\text{cycles of lenght } k\}| = o_k(n).$$

In words, G is locally tree-like.



For example, with probability one, a sequence of uniformly sampled d-regular graphs on n vertices is locally tree-like.

Kesten-McKay Law

Theorem
Fix integer $d \geqslant 2$. Let $G = G_n$ is a sequence of locally tree-like d-regular graphs on n vertices then for any $I \subset \mathbb{R}$,

$$\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}(\lambda_k(X) \in I) = \mu_X(I) \to \mu_{KM}(I).$$

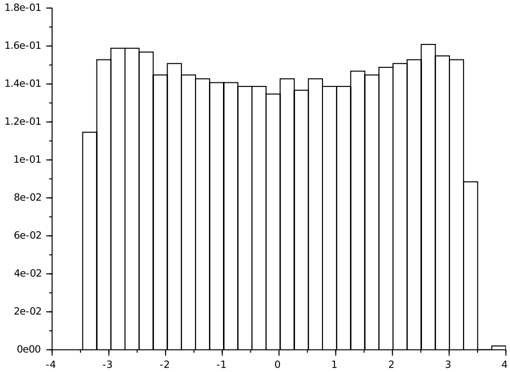
where

$$\mu_{KM}(dx) = \frac{d}{2\pi}\frac{\sqrt{4(d-1)-x^2}}{d^2-x^2}\mathbf{1}_{|x|\leqslant 2\sqrt{d-1}}dx.$$

We have $\mu_{KM}(I\sqrt{d}) \to \mu_{sc}(I)$ when $d \to \infty$.

Simulation

Take $d = 4$, $n = 2000$ and G a uniformly sampled d-regular graph.

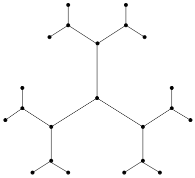


Idea of Proof

Since G is locally tree-like, T is the infinite d-regular tree.

$$\begin{aligned}\int \lambda^k d\mu_X &= \frac{1}{n}\mathrm{Tr}X^k \\ &= \frac{1}{n}\sum_{v=1}^{n} |\{\text{closed walks of length } k \text{ starting from } v \text{ in } G\}| \\ &= |\{\text{closed walks of length } k \text{ starting from the root of } T\}| + o_k(1) \\ &= \int \lambda^k d\mu_{KM} + o_k(1).\end{aligned}$$

It is then fairly easy to compute the number of walks on the infinite regular tree.



Friedman's Theorem

We have $\lambda_1(X) = d$, $\mathrm{supp}(\mu_{KM}) = [-2\sqrt{d-1}, 2\sqrt{d-1}]$ and $\lambda_2(X) \geqslant 2\sqrt{d-1} + o(1)$. Recall that $\lambda_2(X) \leqslant 2\sqrt{d-1}$ for Ramanujan graphs.

Theorem

Fix even integer $d \geqslant 4$. Let $G = G_n$ is a sequence of uniformly distributed d-regular graphs on n vertices, then with high probability,

$$\lambda_2(X) = 2\sqrt{d-1} + o(1) = -\lambda_n(X).$$

Most regular graphs are nearly Ramanujan !

(Friedman 2004)

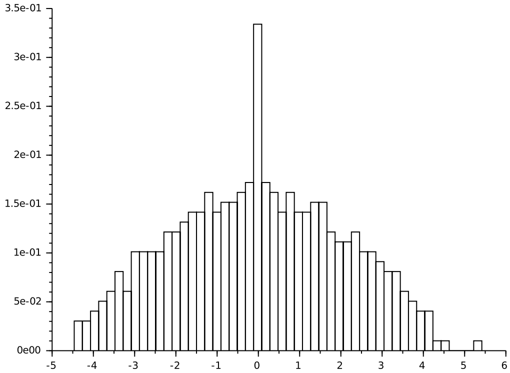
Theorem
Fix integer $c > 0$. *Let* $G = G_n$ *be an Erdős-Rényi graph with parameter* $p = c/n$. *Then, with probability one, for any interval* $I \subset \mathbb{R}$,

$$\mu_X(I) = \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}(\lambda_k(X) \in I) \to \mu_c(I),$$

for some probability measure μ_c *with support* $\mathbb{R}$.

Erdős-Rényi

Histogram of eigenvalues for $c = 4$ and $n = 500$.



The maximum eigenvalue is $\lambda_1(X) = (1 + o(1))\sqrt{\frac{\log n}{\log \log n}}$

(Sudakov and Krivelevich 2003).

There is no explicit expression for μ_c.

Let $\Lambda = \{\lambda_i, i \geqslant 1\}$, be the atoms of μ_c, i.e.

$$\mu_c(\{\lambda\}) > 0 \text{ if and only if } \lambda \in \Lambda.$$

Then Λ is the set algebraic numbers

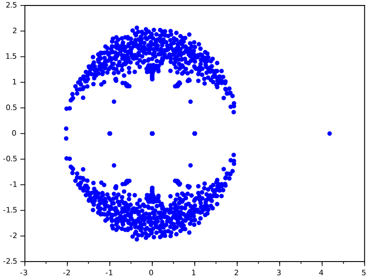
$$\sum_{\lambda \in \Lambda} \mu_c(\{\lambda\}) < 1$$

if and only if $c > 1$.

Also $\mu_c(\{0\})$ has a closed-form expression.

Bordenave/Lelarge/Salez (2012), Salez (2013), Bordenave/Virág/Sen (2014).

Eigenvalues for $c = 4$ and $n = 500$ of the non-backtracking matrix B.



No theorem yet !

In Summary

In Summary

The spectra of well-chosen graph matrices contain a lot of meaningful information and spectral algorithms are reasonably fast.

In Summary

Understanding the spectra of random graphs may notably help to sort meaningful information from noise.

In Summary

As long as the average degree of a vertex grows to infinity, general spectral graph theory, theory of perturbations and random matrices are very useful.

In Summary

The spectrum of diluted random graphs is very far from being understood.

A few general references

Random matrices

G. Anderson, A. Guionnet and O. Zeitouni, *An Introduction to Random Matrices*, 2009.

Z. Bai and J. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, 2006.

Spectral graph theory

L. Lovász, *Eigenvalues of graphs*, 2007.

F. Chung, *Spectral Graph Theory* , 1992.

A Brouwer and W Haemers, *Spectra of Graphs*, 2012.

B. Mohar and W. Woess, *A Survey on Spectra of Infinite Graphs*, 1989.

Markov chains and random walks on graph

D. Aldous and J. Fill, *Reversible Markov Chains and Random Walks on Graphs.*

D. Levin , Y. Peres, E. Wilmer, *Markov Chains and Mixing Times*, 2008.

Thank you for your attention!