Understanding Graphs through Spectral Densities

David Bindel
17 July 2019

Department of Computer Science
Cornell University
Acknowledgements

Thanks Kun Dong and Austin Benson, along with Anna Yesypenko, Moteleolu Onabajo, Jianqiu Wang.

Also: NSF DMS-1620038.
What can we tell from *partial* spectral information (eigenvalues and/or vectors)

Claim: Most spectral analyses involve one of two perspectives:

- Approximate something via a *few* (extreme) eigenvalues.
- Look at *all* the eigenvalues (or all in a range).
“You mean, if you had perfect pitch could you find the shape of a drum.” — Mark Kac (quoting Lipmann Bers)
American Math Monthly, 1966
What information hides in the eigenvalue distribution?

1. Discretizations of Laplacian: something like Weyl’s law
2. Sparse E-R random graphs: Wigner semicircular law
3. Some other random graphs: Wigner semicircle + a bit (Farkas et al, Phys Rev E (64), 2001)
4. “Real” networks: less well understood

Goal: Explore by estimating eigenvalue distributions (fast).
A Bestiary of Matrices

- Adjacency matrix: $A$
- Laplacian matrix: $L = D - A$
- Unsigned Laplacian: $L = D + A$
- Random walk matrix: $P = AD^{-1}$ (or $D^{-1}A$)
- Normalized adjacency: $\bar{A} = D^{-1/2}AD^{-1/2}$
- Normalized Laplacian: $\bar{L} = I - \bar{A} = D^{-1/2}LD^{-1/2}$
- Modularity matrix: $B = A - \frac{dd^T}{2n}$
- Motif adjacency: $W = A^2 \odot A$

All have examples of co-spectral graphs

... through spectrum uniquely identifies quantum graphs
Spectra define a *generalized function* (a density):

$$\text{tr}(f(H)) = \int f(\lambda) \mu(\lambda) \, dx = \sum_{k=1}^{N} f(\lambda_k)$$

where \( f \) is an analytic test function. Smooth to get a picture: a *spectral histogram* or *kernel density estimate*. 
Exploring Spectral Densities

Kernel polynomial method (see Weisse, Rev. Modern Phys.)

- Spectral distribution on $[-1, 1]$ is a generalized function:

$$\int_{-1}^{1} \mu(x)f(x) \, dx = \frac{1}{N} \sum_{k=1}^{N} f(\lambda_k)$$

- Write $f(x) = \sum_{j=1}^{\infty} c_j T_j(x)$ and $\mu(x) = \sum_{j=1}^{\infty} d_j \phi_j(x)$, where

$$\int_{-1}^{1} \phi_j(x) T_k(x) \, dx = \delta_{jk}$$

- Estimate $d_j = \text{tr}(T_j(H))$ by stochastic methods

- Truncate series for $\mu(x)$ and filter (avoid Gibbs)

*Much* cheaper than computing all eigenvalues!

Alternatives: Lanczos (Golub-Meurant), maxent (Röder-Silver)
$Z \in \mathbb{R}^n$ with independent entries, mean 0 and variance 1.

\[
E[(Z \odot HZ)_i] = \sum_j h_{ij} E[Z_iZ_j] = h_{ii}
\]

\[
\text{Var}[(Z \odot HZ)_i] = \sum_j h_{ij}^2.
\]

Serves as the basis for stochastic estimation of

- Trace (Hutchinson, others; review by Toledo and Avron)
- Diagonal (Bekas, Kokiopoulou, and Saad)

Independent probes $\rightarrow$ $1/\sqrt{N}$ convergence (usual MC).
(Can go beyond independent probes.)
Spike (non-smoothness) at eigenvalues of 0 leads to inaccurate approximation.
Motifs and Symmetry

Suppose $PH = HP$. Then

- $\mathcal{V}$ a max invariant subspace for $P$ $\implies$ $\mathcal{V}$ a max invariant subspace for $H$

So local symmetry $\implies$ localized eigenvectors.

Simplest example: $P$ swaps $(i, j)$

- $e_i - e_j$ an eigenvector of $P$ with eigenvalue $-1$
- $e_i - e_j$ an eigenvector of $\overline{A}$ with eigenvalue

$$\lambda = \rho_{\overline{A}}(e_i - e_j) = \begin{cases} 
  d^{-1}, & (i, j) \in \mathcal{E} \\
  0, & \text{otherwise}.
\end{cases}$$

- All other eigenvectors (eigenvalue $-1$) satisfy $v_i = v_j$
Motifs in Spectrum

- $\lambda = 0$
  
- $\lambda = \pm 1/2$

- $\lambda = -1/2$

- $\lambda = \pm 1/\sqrt{2}$
Motif “spikes” slow convergence – deflate motif eigenvectors!

If $P \in \mathbb{R}^{n \times m}$ an orthonormal basis for the quotient space,

- Apply estimator to $P^T \overline{A} P$ to reduce size for $m \ll n$.
- or use $Proj_P(Z)$ to probe the desired subspace.
Diagonal Estimation and LDoS

Diagonal estimation also useful for local DoS $\nu_k(x)$; in the symmetric case with $H = QQ^T$, have

$$\int f(x) \nu_k(x) \, dx = f(H)_{kk} = e_k^T Qf(\Lambda) Q^T e_k$$

$$\nu_k(x) = \sum_{j=1}^{n} q_{kj}^2 \delta(x - \lambda_j)$$

DoS is sum of local densities of states:

$$\mu(x) = \sum_{k=1}^{n} \nu_k(x)$$
KPM for LDoS

Same game, different moments:

- Estimate $d_j = [T_j(H)]_{kk}$ by diag estimation
- Truncate series for $\mu(x)$ and filter (avoid Gibbs)

Diagonal estimator gives moments for all $k$ simultaneously!

Alternatives: Lanczos (Golub-Meurant), maxent (Röder-Silver)
Can compute common *centrality measures* with LDoS

- Estrada centrality: $\exp(\gamma A)_{kk}$
- Resolvent centrality: $[(I - \gamma \bar{A})^{-1}]_{kk}$

Some motifs associated with localized eigenvectors:

- Chief example: Null vectors of $\bar{A}$ supported on leaves.
- Use LDoS + topology to find motifs?

Other uses: clustering and role discovery. What else?
Exploring Spectral Densities (with David Gleich)

- Compute spectrum of normalized Laplacian / RW matrix
- Compare KPM to full eigencomputation

Things we know

- Eigenvalues in $[-1, 1]$; nonsymmetric in general
- Stability: change $d$ edges, have
  \[ \lambda_{j-d} \leq \hat{\lambda}_j \leq \lambda_{j+d} \]
- $k$th moment = $P$ (return after $k$-step random walk)
- Eigenvalue cluster near 1 $\sim$ well-separated clusters
- Eigenvalue cluster near -1 $\sim$ bipartite structure
- Eigenvalue cluster near 0 $\sim$ leaf clusters

What else can we “hear”?
Erdos
Erdos (local)
Internet topology
PGP
PGP (local)
Yeast (local)
$N = 326186, \text{nnz} = 1615400, 80 \text{ s (1000 moments, 10 probes)}$
$N = 1139905, \text{nnz} = 113891327, 2093 \text{ s (1000 moments, 10 probes)}$
Barabási–Albert model

Scale-free network (5000 nodes, 4999 edges)
Small world network (5000 nodes, 260000 edges)
Block Two-Level Erdős-Rényi model (BTER)

- First Phase: Erdős-Rényi Blocks
- Second Phase: Using Chung-Lu Model to connect blocks with $p_{ij} = p(d_i, d_j)$
Model Verification: BTER

Figure 1: Erdos collaboration network.

Figure 2: BTER model for Erdos collaboration network.
Latest:

- Dong, Benson, Bindel (KDD 2019).
- Longer talk at ILAS 2019 (slides online)