Vectorized Laplacians for dealing with high-dimensional data sets

Fan Chung Graham
University of California, San Diego
Vectorized Laplacians ----- Graph gauge theory

- Earlier work (3 papers 1992-1994) with Shlomo Sternberg, Bert Kostant on the vibrational spectra and spins of the molecules.

- New directions in organizing/analyzing high-dimensional data sets, with applications in
  - vector diffusion maps,
  - graphics, imaging, sensing
  - resource allocation with dynamic demands

Singer + Wu 2011, C. + Zhao 2012
Cayley graphs $\Gamma = (V,E)$

$V \leftarrow$ a group $H = A_5$

$u \sim v \iff u = gv$ for $g \in B$, a chosen subset of $H$.

$B = \{(12345),(54321),(12)(34)\} \subset A_5$
Cayley graphs $\Gamma = (V, E)$

$V \leftarrow$ a group $H$

$u \sim v \iff u = gv$ for $g \in B$, a chosen subset of $H$.

adjacency matrix $\approx$

irreducible representations

Frobenius 1897
Eigenvalues of the buckyball
with weight 1 for single bonds and $t$ for double bonds

Roots of

- $(x^2 + x - t^2 + t - 1)(x^3 - tx^2 - x^2 - t^2 x - 3x + t^3 - t^2 + t + 2) = 0$
  with multiplicity 5

- $(x^2 + x - t^2 - 1)(x^2 + x - (t + 1)^2) = 0$
  with multiplicity 4

- $(x^2 + (2t + 1)x + t^2 + t - 1)(x^4 - 3x^3 + (-2t^2 + t - 1)x^2 + (3t^2 - 4t + 8)x + t^4 - t^3 + t^2 + 4t - 4) = 0$
  with multiplicity 3

- $x - t - 2 = 0$
  with multiplicity 1

$5^2 + 4^2 + 3^2 + 3^2 + 1 = 60$
Discrete Laplace operator

\[ \Delta f(x) = \frac{1}{d_x} \sum_{y \sim x} (f(x) - f(y)) \]

For a path \( P_n \)

\[ \Delta f(x_j) = -\frac{1}{2} \left\{ (f(x_{j+1}) - f(x_j)) - (f(x_j) - f(x_{j-1})) \right\} \]

\[ -\frac{\partial^2}{\partial x^2} f(x_j) = -\left\{ \frac{\partial}{\partial x} f(x_{j+1}) - \frac{\partial}{\partial x} f(x_j) \right\} \]
In a graph $\Gamma = (V,E)$, $\Delta$ acts on
\[ \mathcal{F}(V, \mathbb{R}) = \{ f : V \to \mathbb{R} \} \]

In general, we consider
\[ \Delta \text{ acts on } \mathcal{F}(V, X) = \{ f : V \to X \} . \]

Example 1: Vibrational spectra of molecules
\[ \mathcal{F}(V, \mathbb{R}^3) . \]

Example 2: Vector Diffusion Maps
\[ \mathcal{F}(V, \mathbb{R}^d) . \]
Connections: a classical framework for dealing with
- Vibrations and spins of the Buckyball
- Noise in the data sets, etc.
Vector bundle  $\mathcal{E} = \{ \mathcal{E}_x : x \in V \}$
- Vector bundle $\mathcal{E} = \{ \mathcal{E}_x : x \in V \}$
- Connection $T$
Vector bundle $\mathcal{E} = \{ \mathcal{E}_x : x \in V \}$

Connection $T$

$$T_{x,e} : \mathcal{E}_y \rightarrow \mathcal{E}_x \quad \text{for } x \in V, \ e = (x,y) \in E$$

$$T_{x,e} \circ T_{y,e} = id$$

Example 1:

- Vibrational spectra of molecules

  Spins of the Buckyball
Example 1: Vibrational spectra of molecules

\[ \mathcal{F}(V, \mathbb{R}^3) = \{h : V \to \mathbb{R}^3\} \]

the space of displacements
Example 1: Vibrational spectra of molecules

\[ h : V \rightarrow \mathbb{R}^3 \]

Hooke's law:

Potential energy

\[
\begin{align*}
\sum_{u,v} k_{u,v} \left( \| \vec{u} + h(u) - \vec{v} - h(v) \| - \| \vec{u} - \vec{v} \| \right)^2 \\
\simeq \sum_{u,v} k_{u,v} \left\| w_{u,v} \cdot (h(u) - h(v)) \right\|^2 \\
= \sum_{u,v} k_{u,v} < h(u) - h(v), A_e (h(u) - h(v)) > \\
\end{align*}
\]

where \( w_{u,v} = \frac{\vec{u} - \vec{v}}{\| \vec{u} - \vec{v} \|} \), \( A_e = w_{u,v} \otimes w_{u,v}^t \)
The vibrational spectrum (infrared) of the Buckyball
• Vector bundle $\mathcal{E} = \{ \mathcal{E}_x : x \in V \}$

• Connection $T$

\[ T_{x,e} : \mathcal{E}_y \to \mathcal{E}_x \quad \text{for } x \in V, \]

\[ T_{x,e} \circ T_{y,e} = id \quad \text{for } e = (x,y) \in E \]

• Selection $\text{sec}\mathcal{E}$

\[ s(u) \in \mathcal{E}_u \quad \text{for } s \in \text{sec}\mathcal{E} \]

Suppose a group $G$ acts on $V$ as graph automorphisms.

For $a \in G$ and $u \in V$, $a : \mathcal{E}_u \to \mathcal{E}_{au}$

$$(a \cdot T)_{x,e} = aT_{a^{-1}x,a^{-1}e}a^{-1}$$

Define $Q_{T,A}(s) = \sum_{x,e} \left( s(x) - T_{x,e}s(y) \right) \cdot A_{x,e} \left( s(x) - T_{x,e}s(y) \right)$
Example 1: Vibrations and spins on the Buckyball

Theorem: C.+Sternberg 1992

In a homogenous graph $\Gamma = \mathcal{G}/\mathcal{H}$ with edge generating set $B$, a $\mathcal{G}$-invariant connection on $\mathcal{E} = \mathcal{E}(\Gamma, X)$ defines an isomorphism $t_b \in Auto(X)$, such that the operator

$$Q_{T,A}(s) = \sum_{x,e} (s(x) - T_{x,e}s(y)) \cdot A_{x,e} (s(x) - T_{x,e}s(y))$$

can be represented as

$$F = \left( \sum_{b \in B} A_{v,be} \right) \otimes I - \sum_{b \in B} A_{v,be} t_b \otimes r(b)$$

For the Buckyball $\mathcal{G} = SL(2,5)$, $PSL(2,5) \sim A_5$,

Isotropy subgroup $\mathbb{Z}_2$
Examples of connection graphs

Example 1: Vibrational spectra of molecules

Spins of the buckyball

Example 2: Vector Diffusion Maps  

\[ O_{uv} \]

rotation

data point \( u \)

\[ O_{uv} \]

data point \( v \)
Examples of connection Laplacians

Example 2: Vector Diffusion Maps

At a node $u$, data $\rightarrow$ local PCA dimension reduction $\rightarrow$ vector in $\mathbb{R}^d$

In a network, rotation $O_{uv}$

Saturday, August 11, 2012
Consistency of connection Laplacians

For a connected connection graph $G$ of dimension $d$,

For $f : V \rightarrow \mathbb{R}^d$

$$< f , L f > = \sum_{(i,j) \in E} (|f(u)O_{uv} - f(v)|^2 w_{uv})$$
Consistency of connection Laplacians

For a connected connection graph $G$ of dimension $d$, the following statements are equivalent:

- The $0$ eigenvalue has multiplicity $d$.
- All eigenvalues have multiplicity $d$.
- For any two vertices $u$ and $v$, any two paths $P, P'$,
  $$\prod_{xy \in P} O_{xy} = \prod_{zw \in P'} O_{zw}$$
- For each vertex $v$, we have $O_v \in SO(d)$, such that
  $$O_{uv} = O_u^{-1} O_v$$ for all $(u, v) \in E$. 

Saturday, August 11, 2012
Consistency of connection Laplacians

Theorem:
For a connected connection graph $G$ of dimension $d$,
if the 0 eigenvalue has multiplicity $d$, then
we can find $O_v \in SO(d)$, such that $O_{uv} = O_u^{-1}O_v$
for all $(u,v) \in E$.

Method 1: For a fixed vertex $u$, choose a frame $O_u$
consisting of any $d$ orthonormal vectors.
For any $v$, define $O_v = \prod O_{v_iv_{i+1}}$ where $u = v_0, v_1, ..., v_t$
is any path joining $u$ and $v$.

Method 2: Use eigenvectors. Singer+Spielman 2012
Dealing with noise in the data

Several combinatorial approaches

- Random connection Laplacians
- Graph sparsification algorithms
- PageRank algorithms
- Graph limits and graphlets

.........
Dealing with noise in the data
Several combinatorial approaches

- Random connection Laplacians
- Graph sparsification algorithms
- PageRank algorithms
- Graph limits and graphlets

........
A random graph model for any given expected degree sequence

Independent random indicator variables

\[ X_{ij} = X_{ji}, 1 \leq i \leq j \leq n. \]

\[ \Pr(X_{ij} = 1) = \Pr(v_i \sim v_j) = p_{ij} \]

\[ X = \begin{pmatrix} X_{ij} \end{pmatrix} \]

Example: The Erdős-Renyi model \( G(n, p) \) with \( p_{ij} \) all equal to \( p \).
A matrix concentration inequality

Theorem: C. + Radcliffe 2011

\[ X_i : \text{independent random } n \times n \text{ Hermitian matrix} \]

\[ X_i - E(X_i) \leq M \]

Then \[ X = \sum_{i=1}^{m} X_i \] satisfies

\[ \Pr\left( \|X - E(X)\| > a \right) \leq 2n \exp\left( \frac{-a^2}{2v^2 + 2Ma/3} \right) \]

where \[ v^2 = \left\| \sum_{i=1}^{m} \text{var}(X_i) \right\| \]
Eigenvalues of random graphs with general distributions

Theorem: C. + Radcliffe 2011

\[ G : \text{ a random graph, where } \Pr(v_i \sim v_j) = p_{ij}. \]

\[ A : \text{ adjacency matrix of } G \]

\[ \bar{A} : \text{ expected adjacency matrix } \bar{A}_{ij} = p_{ij} \]

With probability \( 1 - \varepsilon \), eigenvalues of \( A \) and \( \bar{A} \) satisfy

\[
\left| \lambda_i(A) - \lambda_i(\bar{A}) \right| \leq \sqrt{4d_{\text{max}} \log(2n / \varepsilon)}
\]

for all \( i, 1 \leq i \leq n \), if the maximum degree satisfies

\[
d_{\text{max}} > \frac{4}{9} \log(\frac{2n}{\varepsilon}).
\]
Eigenvalues of the normalized Laplacian for random graphs with general distributions

Theorem: C. + Radcliffe 2011

$G$: a random graph, where $\Pr(v_i \sim v_j) = p_{ij}$.

$A$: adjacency matrix of $G$

$\bar{A}$: expected adjacency matrix $\bar{A}_{ij} = p_{ij}$

$D$: diagonal degree matrix of $G$

With probability $1 - \varepsilon$, eigenvalues of $L = I - D^{-1/2} AD^{-1/2}$ satisfy

$$\left| \lambda_i (L) - \lambda_i (\bar{L}) \right| \leq \sqrt{\frac{3 \log(4n / \varepsilon)}{\delta}}$$

for all $i, 1 \leq i \leq n$, and $\bar{L} = I - \bar{D}^{-1/2} \bar{A} \bar{D}^{-1/2}$, if the minimum degree satisfies $\delta > 10 \log n$. 
Dealing with noise in the data
Several combinatorial approaches

- Random connection Laplacians

- **Graph sparsification algorithms**

- PageRank algorithms

- **Graph limits and graphlets**

- ........
Graph sparsification

For a graph $G=(V,E,W)$,

we say $H$ is an $\varepsilon$-sparsifier if for all $S \subseteq V$ we have

$$|h_G(S) - h_H(S)| \leq \varepsilon h_G(S).$$

Benczur and Karger 1994 --- sparsifier with size $O(n \log n)$

in running time $\tilde{O}(n^2)$

Spielman and Sriastava 2008 --- using effective resistance

running time $\tilde{O}(n)$

Batson, Spielman and Sriastava 2009 --- $O(n)$ sparsifier

running time $O(nm)$
Graph sparsification

\[ |h_G(S) - h_H(S)| \leq \varepsilon h_G(S) \]

The cut edges are more likely to be sampled in order to satisfy, for all \( S \subseteq V \),

We need to rank edges.
Dealing with noise in the data
Several combinatorial approaches

- Random connection Laplacians
- Graph sparsification algorithms
- \textbf{PageRank algorithms for ranking edges}
- Graph limits and graphlets
- ........
PageRank for ranking vertices

Two equivalent ways to define PageRank $p = pr(\alpha, s)$

(1) $p = \alpha s + (1 - \alpha) p W$

(2) $p = \alpha \sum_{t=0}^{\infty} (1 - \alpha)^t (s W^t)$

$s = \left( \frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n} \right)$  \[\rightarrow\]  the (original) PageRank

$s = \text{some "seed"}, \text{ e.g., } (1, 0, \ldots, 0)$  \[\rightarrow\]  personalized PageRank
PageRank as generalized Green’s function

Discrete Green’s function

Laplacian
\[ \mathcal{L} = D^{1/2} \Delta D^{-1/2} = \sum_{k=1}^{n-1} \lambda_k P_k \]

Green’s function
\[ \mathcal{G} = \sum_{k=1}^{n-1} \frac{1}{\lambda_k} P_k \]
\[ \mathcal{L} \mathcal{G} = I \quad \text{restricted to } (\ker \mathcal{L})^\perp \]

The general Green’s function
\[ \mathcal{G}_\beta = \sum_{k=1}^{n-1} \frac{1}{\beta + \lambda_k} P_k \]

PageRank
\[ pr(\alpha, s) = \beta s D^{-1/2} \mathcal{G}_\beta D^{1/2}, \quad \beta = \frac{2\alpha}{\sqrt{1-\alpha}} \]
Ranking pages using Green’s function and electrical network theory

\[ G = (V, E, W) \] : a weighted graph

\[ i_v : V \rightarrow \mathbb{R} \] injected current function

\[ i_E : E \rightarrow \mathbb{R} \] induced current function

\[ B(e, v) = \begin{cases} 
1 & \text{if } v \text{ is } e's \text{ head} \\
-1 & \text{if } v \text{ is } e's \text{ tail} \\
0 & \text{otherwise}
\end{cases} \]

Kirchoff’s current law: \[ i_v = i_E B \]

Ohm’s current law: \[ i_E = f B^T W \]

\[ L = B^T W B \]
Ranking pages using Green’s function and electrical network theory

Kirchoff’s current law: \[ i_V = i_E B \]

Ohm’s current law: \[ i_E = f \ B^T W \]

\[ i_V = f_V B^T W_{\beta} B = f_V L_{\beta} \]

\[ f_V = i_V D^{-1/2} \mathcal{G}_{\beta} D^{-1/2} \]

\[ R_{\beta}(u,v) = f_V (\chi_u - \chi_v)^T \]

\[ = i_V D^{-1/2} \mathcal{G}_{\beta} D^{-1/2} (\chi_u - \chi_v)^T \]

\[ = (\chi_u - \chi_v) D^{-1/2} \mathcal{G}_{\beta} D^{-1/2} (\chi_u - \chi_v)^T \]

\[ pr_{\alpha}(s) = sD^{-1/2} \mathcal{G}_{\beta} D^{1/2} \quad \text{where} \quad \beta = \frac{2\alpha}{1 - \alpha} \]
Ranking pages using Green’s function and electrical network theory

The Green value of \((u,v)\)

\[ R_\alpha(u,v) = h_\alpha(u,v) + h_\alpha(v,u) \]

\[ = \frac{pr_{\alpha,u}(u)}{d_u} - \frac{pr_{\alpha,u}(v)}{d_v} + \frac{pr_{\alpha,v}(v)}{d_v} - \frac{pr_{\alpha,v}(u)}{d_u} \]
Vectorized Laplacian and vectorized PageRank

A connection graph
\[ \mathcal{G} = (V,E,O,w) \]

A connection matrix
\[ A = \begin{cases} w_{uv} O_{uv} & \text{if } (u,v) \in E, \\ 0_{d \times d} & \text{otherwise}. \end{cases} \]

\[ \mathcal{D}(v,v) = d_u I_{d \times d} \]

\[ \mathbb{L} = \mathcal{D} - A \]

\[ f^T \mathbb{L} f = \sum_{(u,v) \in E} w_{uv} \| f(u)O_{uv} - f(v) \|^2 \]
Vectorized Laplacian and vectorized PageRank

A connection graph \( G = (V, E, O, w) \)

The transition probability matrix

\[
P = D^{-1} A
\]

\[
Z = \frac{I + P}{2}
\]

The vectorized PageRank

\[
\hat{p}(\alpha, v) = \alpha \hat{s} + (1 - \alpha) \hat{p}Z
\]
Ranking pages using Green's function and electrical network theory

The Green value of \((u, v)\)

\[
R_\alpha(u, v) = \left\| \frac{\hat{p}_{\alpha,u}(u)}{d_u} - \frac{\hat{p}_{\alpha,u}(v)}{d_v} + \frac{\hat{p}_{\alpha,v}(v)}{d_v} - \frac{\hat{p}_{\alpha,v}(u)}{d_u} \right\|
\]

which the sampling subroutine is based upon.
Dealing with noise in the data
Several combinatorial approaches

- Random connection Laplacians
- Graph sparsification algorithms
- PageRank algorithms for ranking edges
- Graph limits and graphlets
- ........
Examples of connection Laplacians

Example 1: Vibrational spectra of molecules
    Spins of the buckyball

Example 2: Electron microscopy
    Vector Diffusion Maps

Example 3: Graphics, vision, ...
    Managing millions of images, photos.
Millions of photos in a box!

Navigate using connection graphs.
- finding consistent subgraphs, paths ...
Feature detection

Detect features using SIFT [Lowe, IJCV 2004]
Feature matching

Match features between each pair of images
Agarwal, Snavely, Simon, Seitz, Szeliski 2009

Part of an image connection graph

Building Rome in a day
Vectorized Laplacians for dealing with high-dimensional data sets

Fan Chung Graham
University of California, San Diego
Thank you.
Part of an image connection graph

Building Rome in a day

Agarwal, Snavely, Simon, Seitz, Szeliski 2009
An example of connection graphs
Building Rome in a day
Agarwal, Snavely, Simon, Seitz, Szeliski 2009