Two questions about randomized in numerical linear algebra

1. Multiway random walks for reducing variance (arxiv.org/1608.04361)
2. An eigenvalue law for “triangle Laplacians”
3. An approximation bound for sampling zonotopes for clusters.

David F. Gleich
Purdue University
3 slides to pose a simple question.
I don’t have an answer 😊
Eigenvalue laws of random ER graphs are well known

The adjacency matrix has a semi-circle law.

The normalized Laplacian matrix also has a semi-circle law with a specific gap

A idea based on our recent work

given a graph $G = (V, E)$ and its adjacency matrix $A$

consider using the weighted matrix $W = A^2 \odot A$

The matrix $W = A^2 \odot A$ arises from our motif and higher-order clustering framework when using triangles as the motif.
The eigenvalues of the normalized Laplacian of this weighted matrix do not follow a Marchenko–Pastur law

**Method.**

Generate a random ER graph, compute

\[ W = A^2 \odot A \]

Compute eigenvalues of the normalized Laplacian of the largest component of \( W \)

**Our question.**

What is this distribution?

Code to reproduce https://gist.github.com/dgleich/4d4becc858e4a7d7952af6c66c99e7b9
Another problem where I don’t know the solution.


We have a heuristic randomized (and useful) algorithm that we don’t have good analysis for 😊
Correlation clustering involves a weighted, signed graph.

Edges in a signed graph indicate similarity (+) or dissimilarity (-).
Correlation clustering involves a weighted, signed graph.

Edges in a signed graph indicate similarity (+) or dissimilarity (-).

Edges can be weighted, but problems become harder.
Edges in a signed graph indicate similarity (+) or dissimilarity (-).

Objective: Minimize the weight of “mistakes”
Edge weights can be stored in an adjacency matrix

\[ A_{ij} = w_{ij}^+ - w_{ij}^- \]

\[
A = \begin{bmatrix}
0 & -6 & +2 & -4 \\
-6 & 0 & -3 & +6 \\
+2 & -3 & 0 & -2 \\
-4 & +6 & -2 & 0 \\
\end{bmatrix}
\]
The rank-1 positive semidefinite case is very simple

\[ A = vv^T \]

\[
A = \begin{bmatrix}
-2 & +3 & -1 & +2 \\
+3 & -2 & +3 & -1 \\
-1 & +3 & -2 & +3 \\
+2 & -1 & +3 & -2
\end{bmatrix}
\]
The rank-1 positive semidefinite case is very simple

\[ A = \mathbf{v}\mathbf{v}^\top \]

\[
A = \begin{bmatrix}
-2 & +3 & -1 & +2 \\
+3 & -2 & -1 & +2 \\
-1 & -1 & -2 & +3 \\
+2 & +2 & +3 & -2 \\
\end{bmatrix}
\]
The rank-1 positive semidefinite case is very simple

\[ A = vv^T \]

\[ A = \begin{bmatrix} -2 & +3 & -1 & +2 \\ +3 & -1 & +2 \end{bmatrix} \]
The rank-1 positive semidefinite case is very simple

Ordering \( \mathbf{v} \) gives a perfect clustering

\[ \mathbf{A} = \mathbf{vv}^\top \]

\[
\begin{bmatrix}
-2 \\
+3 \\
-1 \\
+2
\end{bmatrix}
\rightarrow
\begin{bmatrix}
-2 \\
-1 \\
+2 \\
+3
\end{bmatrix}
\]
The rank-1 positive semidefinite case is very simple

\[ A = \mathbf{v}\mathbf{v}^T \]

Ordering \( \mathbf{v} \) gives a perfect clustering
The rank-1 positive semidefinite case is very simple.

Ordering $\mathbf{v}$ gives a perfect clustering.

\[ A = \mathbf{v} \mathbf{v}^T \]

\[
\begin{pmatrix}
-2 \\
+3 \\
-1 \\
+2
\end{pmatrix}
\rightarrow
\begin{pmatrix}
-2 \\
-1 \\
+2 \\
+3
\end{pmatrix}
What happens for other low-rank matrices?

A simple solution for rank-1 positive semidefinite correlation clustering always exists.

Veldt, Gleich, Wirth arXiv:1611.07305

• A single negative eigenvalue makes the problem NP-hard 😞
• A rank-\(r\) pos def matrix can be solved in “polynomial” time \(O(n^{r^2-r})\).
  Rank 3 = \(O(n^8)\) time
• Algorithm is based on equiv between low-rank CC and vector partitioning
• In Stinson, Gleich, Constantine (arXiv:1602.06620) we proposed a randomized algorithm to find vertices of a zonotope.
Vector partitioning formulation

$max \sum_{k=1}^{\text{# clusters}} \| S_k \|^2$

Cluster $C_i$

“Sum point” $S_i$
A zonotope is the linear projection of a hypercube into a lower dimension.

\[ \mathcal{Z}(G) = \text{conv}\{Gx \mid x \in \{\pm 1\}^n\} \]

\[ G \in \mathbb{R}^{D \times N} \quad \text{“generator” matrix} \]
A really simple randomized algorithm to get a vertex of the zonotope

To generate a non-random vertex of a zonotope, compute

$$G \text{sign}(G^T x)$$ where $$x_i \in N(0, 1)$$
Prob. that a vertex is generated depends on normal cone

$N_Z(v)$
Vertices of the signing zonotope correspond to clusterings

1. Begin with “signing” vector

\[ \sigma = (\sigma_{r,s}^i) \in \{\pm 1\}^{n\binom{d+1}{2}} \]

2. Map signings into a zonotope
Vertices of the signing zonotope correspond to clusterings

1. Begin with “signing” vector

\[ \mathbf{G}_\sigma \]

\[ \sigma = (\sigma_{r,s}^i) \in \{\pm 1\}^{n \binom{d+1}{2}} \]

2. Map signings into a zonotope

Columns of \( \mathbf{G} \) are vectorized outer products:

\[ \mathbf{v}_i \cdot (\mathbf{e}_r - \mathbf{e}_s)^T \]
Vertices of the signing zonotope correspond to clusterings

1. Begin with “signing” vector
   \[ \sigma = (\sigma^i_{r,s}) \in \{\pm 1\}^{n(d+1)} \]
2. Map signings into a zonotope
3. If \( \sigma \) is mapped to a zonotope vertex, it defines a clustering

Columns of \( \mathbf{G} \) are vectorized outer products:

\[ \mathbf{v}_i \cdot (\mathbf{e}_r - \mathbf{e}_s)^\top \]
Vertices of the signing zonotope correspond to clusterings

1. Begin with “signing” vector

\[ \sigma = (\sigma_{r,s}^i) \in \{\pm 1\}^{n(d+1)} \]

2. Map signings into a zonotope

3. If \( \sigma \) is mapped to a zonotope vertex, it defines a clustering

\[ \sigma_{r,s}^i = -1 \]

“Node i is not in cluster r, but could be in cluster s”
Vertices of the signing zonotope correspond to clusterings

1. Begin with “signing” vector
\[ \sigma = (\sigma_{r,s}^i) \in \{\pm 1\}^{n(d+1)} \]
2. Map signings into a zonotope
3. If \( \sigma \) is mapped to a zonotope vertex, it defines a clustering
4. There are \( O(n^{d^2-1}) \) vertices

\[ \sigma_{r,s}^i = -1 \]

“Node \( i \) is not in cluster \( r \), but could be in cluster \( s \)”
In practice we just sample vertices of the zonotope

**ZonoCC Algorithm O(nk)**

1. Generate $k$ random extremal signings $\sigma$ via random sampling
2. Find the clustering of each $\sigma$
3. Output clustering with highest objective score

![Graph of Weighted Agreements vs Number of Iterations $k$](image)

$n = 3000, d = 5$, synthetic.
For a rank-3 problem on clustering conferences

Table 2: Objective scores and runtimes in seconds for correlation clustering algorithms on two real-world datasets. Due to the size of the stocks dataset, we can run only ZonoCC and Pivot on it.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>ZonoCC</th>
<th>Pivot</th>
<th>CGW</th>
<th>ILP</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS Conf. Obj.</td>
<td>7540.0</td>
<td>7540.0</td>
<td>7540.0</td>
<td>7540.0</td>
</tr>
<tr>
<td>n = 157 Time</td>
<td>7</td>
<td>1</td>
<td>1380</td>
<td>52</td>
</tr>
<tr>
<td>Stocks Obj.</td>
<td>5100.2</td>
<td>5099.5</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>n = 497 Time</td>
<td>40</td>
<td>20</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>
Back to the question

Are there better ways of sampling these zonotopes?

\[ G \text{sign}(G^T x) \text{ where } x_i \in N(0, 1) \]

Is there anything like an approximation bound we can give for an algorithm based on sampling these in terms of the objective?

- [Ferrez, Fukuda, Libeling] reduce low-rank binary QP maximization to zonotope enumeration

\[ \max x^T A x \text{ such that } x_i \in \{0, 1\} \]
Summary. Here are my two questions!

What is this distribution??

Can we bound the approximation of the random zonotope sampling algorithm?

$$\text{sign}(\mathbf{G}^T \mathbf{x})$$ where $$x_i \in N(0, 1)$$

Code to reproduce https://gist.github.com/dgleich/4d4becc858e4a7d7952af6c66c99e7b9

David Gleich · Purdue