A new discrepancy definition for hypergraphs

Steve Butler

1 Discrepancy on simple graphs

For an undirected graph we let A denote the adjacency matrix (i.e., given vertices u and v we have that $A_{uv} = 1$ if u and v are adjacent and 0 otherwise), and D the diagonal degree matrix. The normalized adjacency matrix is given by $D^{-1/2}AD^{-1/2}$. We let ψ_X denote the indicator vector for a subset X of the vertices. The volume of X, denoted $\operatorname{vol} X$, is $\operatorname{vol} X = \sum_{u\in X} d_u = \|D^{1/2}\psi_X\|^2$.

The discrepancy of a graph provides a measurement on how randomly the edges are distributed. We define the discrepancy of a graph G, denoted $\operatorname{disc} G$, to be the minimal β such that

$$\Big|e(X,Y) - \frac{\operatorname{vol} X\ \operatorname{vol} Y}{\operatorname{vol} G}\Big| \le \beta\sqrt{\operatorname{vol} X\ \operatorname{vol} Y}.$$

Here $e(X,Y)$ denotes the number of edges joining a vertex in X to a vertex in Y.

It is known that

$$\operatorname{disc} G \le \sigma_2(D^{-1/2}AD^{-1/2}) \le 18\operatorname{disc} G\big(1 - \frac{5}{2}\log\operatorname{disc} G\big),$$

where σ_2 denotes the second largest singular value (which for symmetric matrices corresponds to the second largest normed eigenvalue). This shows that $\operatorname{disc} G$ and $\sigma_2(D^{-1/2}AD^{-1/2})$ are equivalent in that if one goes to zero then both go to zero.

2 Generalizing to hypergraphs

The difficulty with generalizing discrepancy to hypergraphs is how to handle “adjacent”. For a k-graph one approach is to replace the matrix with a multi-dimensional matrix (i.e., a k-dimensional array) with 1 in an entry if and only if the union of the indices corresponds to an edge. An obvious difficulty with this is the matrix is *very* sparse, but even worse there is no generalization of eigenvalues and/or eigenvectors. The approach we take here will be different.

For a k-graph we fix an i with $0 < i < k$. We then will define an adjacency matrix $A^{(i,k-i)}$ by indexing the rows by the i element subsets of V, indexing the columns by the $k-i$ element subsets of V, and an entry of $A^{(i,k-i)}$ is 1 if the union of the sets indexing the row and column gives a k-edge and 0 otherwise. Note that for a 2-graph that $A^{(1,1)}$ gives the adjacency matrix defined above. (We note that in general this matrix will not be symmetric, so we will need to work with singular values instead of eigenvalues.)

Given an i element subset X of V we define the analogue of the degree by

$$\Delta^{(i)}(X) = \left| \left\{ Y : \begin{array}{c} Y \text{ a } k-i \text{ element subset of V} \\ \text{and } X \cup Y \text{ is a } k\text{-edge of } G \end{array} \right\} \right|$$

Note that $\Delta^{(i)}(X)$ is the row sum of $A^{(i,k-i)}$ which corresponds to X, while $\Delta^{(k-i)}(Y)$ is the column sum of $A^{(i,k-i)}$ which corresponds to Y. We also have the analogues of the diagonal degree matrix, $\Delta^{(i)}$ and $\Delta^{(k-i)}$. This gives us our analogue of the normalized adjacency matrix, namely,

$$(\Delta^{(i)})^{-1/2} A^{(i,k-i)} (\Delta^{(k-i)})^{-1/2}.$$

Given $\mathcal{X}^{(i)}$ and $\mathcal{Y}^{(k-i)}$, collections of i and $k-i$ element subsets respectively, then we have that the number of k-edges formed by these subsets is given by

$$e(\mathcal{X}^{(i)}, \mathcal{Y}^{(k-i)}) = \left| \left\{ (X,Y) : X \in \mathcal{X}^{(i)}, Y \in \mathcal{Y}^{(k-i)}, \text{ and } X \cup Y \text{ a k edge} \right\} \right|.$$

We also need the other half of discrepancy, namely, the "expected" number of k-edges given a collection of i and $k-i$ element subsets. Let us suppose that $\mathcal{X}^{(i)}$ is a collection of i element subsets of the vertices, we define the volume of $\mathcal{X}^{(i)}$ analogously as before by

$$\mathrm{vol}^{(i)}(\mathcal{X}^{(i)}) = \sum_{X \in \mathcal{X}^{(i)}} \Delta^{(i)}(X),$$

and similarly if we have $\mathcal{Y}^{(k-i)}$ a collection of $k-i$ element subsets of the vertices, we have

$$\mathrm{vol}^{(k-i)}(\mathcal{Y}^{(k-i)}) = \sum_{Y \in \mathcal{Y}^{(k-i)}} \Delta^{(k-i)}(Y).$$

We note in passing that $\mathrm{vol}^{(i)}(G) = \mathrm{vol}^{(k-i)}(G) := \mathrm{vol}^{(i,k-i)}(G)$ denotes the total number of nonzero entries in $A^{(i,k-i)}$.

We are now ready to define the $(i,k-i)$-discrepancy of G, denoted $\mathrm{disc}^{(i,k-i)}(G)$. It is the minimum β such that

$$\left| e(\mathcal{X}^{(i)}, \mathcal{Y}^{(k-i)}) - \frac{\mathrm{vol}^{(i)}(\mathcal{X}^{(i)})\ \mathrm{vol}^{(k-i)}(\mathcal{Y}^{(k-i)})}{\mathrm{vol}^{(i,k-i)}(G)} \right| \le \beta \sqrt{\mathrm{vol}^{(i)}(\mathcal{X}^{(i)})\ \mathrm{vol}^{(k-i)}(\mathcal{Y}^{(k-i)})}.$$

Similar to before we have

$$\mathrm{disc}^{(i,k-i)}(G) \le \sigma_2\big((\Delta^{(i)})^{-1/2}A^{(i,k-i)}(\Delta^{(k-i)})^{-1/2}\big) \le 150\,\mathrm{disc}^{(i,k-i)}(G)(1-8\log\mathrm{disc}^{(i,k-i)}(G)).$$

We note that $\sigma_2\big((\Delta^{(i)})^{-1/2}A^{(i,k-i)}(\Delta^{(k-i)})^{-1/2}\big) \le 1$. In particular, we see that again both of these properties are equivalent in that if one goes to zero then both go to zero.

3 Quick sketch of key ideas

The singular values of $(\Delta^{(i)})^{-1/2}A^{(i,k-i)}(\Delta^{(k-i)})^{-1/2}$ are the square root of the eigenvalues of

$$\begin{aligned}\big((\Delta^{(i)})^{-1/2}A^{(i,k-i)}(\Delta^{(k-i)})^{-1/2}\big)^*\big((\Delta^{(i)})^{-1/2}A^{(i,k-i)}(\Delta^{(k-i)})^{-1/2}\big)&\\ = (\Delta^{(k-i)})^{-1/2}A^{(k-i,i)}(\Delta^{(i)})^{-1}A^{(i,k-i)}(\Delta^{(k-i)})^{-1/2}).&\end{aligned}$$

It is straightforward to check that $(\Delta^{(k-i)})^{1/2}\mathbf{1}$ is an eigenvector of this matrix with eigenvalue 1. Since this matrix is nonnegative and the eigenvector positive it follows that 1 is the largest singular value with vectors $(\Delta^{(i)})^{1/2}\mathbf{1}$ (on the left) and $(\Delta^{(k-i)})^{1/2}\mathbf{1}$ (on the right). It follows that

$$\begin{aligned}\sigma_1\Big((\Delta^{(i)})^{-1/2}A^{(i,k-i)}(\Delta^{(k-i)})^{-1/2} - \frac{(\Delta^{(i)})^{1/2}J(\Delta^{(k-i)})^{1/2}}{\mathrm{vol}^{(i,k-i)}\,G}\Big)&\\ = \sigma_2\big((\Delta^{(i)})^{-1/2}A^{(i,k-i)}(\Delta^{(k-i)})^{-1/2}\big),&\end{aligned}$$

where J is the all 1's matrix.

Now we note that

$$\begin{aligned}\Big|e(\mathcal{X}^{(i)},\mathcal{Y}^{(k-i)}) - \frac{\mathrm{vol}^{(i)}(\mathcal{X}^{(i)})\ \mathrm{vol}^{(k-i)}(\mathcal{Y}^{(k-i)})}{\mathrm{vol}^{(i,k-i)}(G)}\Big| &= \Big|\Big\langle\psi_{\mathcal{X}^{(i)}}, A^{(i,k-i)} - \frac{\Delta^{(i)}J\Delta^{(k-i)}}{\mathrm{vol}^{(i,k-i)}\,G}\psi_{\mathcal{Y}^{(k-i)}}\Big|\Big\rangle\\ = \Big|\Big\langle(\Delta^{(i)})^{1/2}\psi_{\mathcal{X}^{(i)}}, \Big((\Delta^{(i)})^{-1/2}A^{(i,k-i)}(\Delta^{(k-i)})^{-1/2} - \frac{(\Delta^{(i)})^{1/2}J(\Delta^{(k-i)})^{1/2}}{\mathrm{vol}^{(i,k-i)}\,G}\Big)&(\Delta^{(k-i)})^{-1/2}\psi_{\mathcal{Y}^{(k-i)}}\Big|\Big\rangle\\ \le \sigma_2\big((\Delta^{(i)})^{-1/2}A^{(i,k-i)}(\Delta^{(k-i)})^{-1/2}\big)\big\|(\Delta^{(i)})^{1/2}\psi_{\mathcal{X}^{(i)}}\big\|\ \big\|(\Delta^{(k-i)})^{-1/2}\psi_{\mathcal{Y}^{(k-i)}}\big\|&\\ = \sigma_2\big((\Delta^{(i)})^{-1/2}A^{(i,k-i)}(\Delta^{(k-i)})^{-1/2}\big)\sqrt{\mathrm{vol}^{(i)}(\mathcal{X}^{(i)})\ \mathrm{vol}^{(k-i)}(\mathcal{Y}^{(k-i)})}.&\end{aligned}$$

Here $\psi_{\mathcal{X}^{(i)}}$ and $\psi_{\mathcal{Y}^{(k-i)}}$ denotes the indicator functions for $\mathcal{X}^{(i)}$ and $\mathcal{Y}^{(k-i)}$ respectively.

The proof for the other bound is done in a note for the discrepancy for directed graphs.