Lecture Notes on an Upper Bound for the Best Cut Quotient from a Vector

Stephen Guattery

Revised 7 May 1996

1 Notation

The following notation conventions are used in these notes:

- Capital letters represent matrices and bold lower-case letters represent vectors. For a matrix A, a_{ij} denotes the element in row i and column j; for the vector $\mathbf{x}$, x_i denotes the i^{th} entry in the vector.
- Various special matrices are represented by the following conventions: The adjacency matrix is denoted Adj; the degree matrix is denoted D; the Laplacian $D - Adj$ is denoted A. The Laplacian is sometimes referred to as the difference Laplacian; the "sum Laplacian" will be the matrix $D + Adj = 2\,D - A$, which will be denoted as B.
- The notion of Laplacian can be extended to graphs with positive edge weights. In particular, let edge (i, j) have weight w_{ij}. The adjacency matrix is modified so that entry $Adj_{ij} = w_{ij}$. The degree of a vertex is defined as the sum of the weights of the incident edges. The definitions for D, A, and B are as above with respect to these changes. Following Fiedler, we will refer to A in the weighted case as the **generalized Laplacian**. We will refer to B in the weighted case as the generalized sum Laplacian. The Laplacian can be considered as the generalized Laplacian where all edge weights are 1.
- The vector that has all entries equal to one is denoted as $\vec{1}$.
- Δ represents the maximum degree of a graph. If the graph has weighted edges, the generalized definition of degree given above applies.
- Let S denote the set of edges forming an edge separator that separates vertex sets V_1 and V_2. Then
$$q(S) = \frac{|S|}{\min(|V_1|, |V_2|)}$$
is called the **cut quotient** for S. If the graph has positive edge weights, the size of the cut is replaced by the total weight of the cut in the definition above.

- A vector $\mathbf{x}$ can be thought of as assigning values to the vertices of a graph G. Assume $\mathbf{x}$ has $k > 1$ distinct values $t_1 < t_2 < \ldots < t_k$, and consider any cut that separates the vertices with values less than or equal to t_i ($i < k$) from those with greater values. Such a cut is called a **threshold cut** based on $\mathbf{x}$.

2 Background Notes

The proof below was formulated by Steve Guattery and Gary Miller. It is a different proof of a result from Spielman and Teng's paper *Spectral Partitioning Works: Planar Graphs and Finite Element Meshes*, which is currently available as a preprint.

This proof is a generalization of Mohar's proof from *Isoperimetric Numbers of Graphs* (Journal of Combinatorial Theory, Series B v.47, pp 274–291 (1989)). In particular, the proof has been extended to apply to vectors other than the second eigenvector of the Laplacian at the cost of loosening the bound slightly for certain vectors. It also applies to graphs with positive edge weights.

3 The Proof

Theorem 3.1 *Let G be a connected graph with positive edge weights on n vertices with generalized Laplacian A. For any vector $\mathbf{x}$ such that $\mathbf{x}^T\vec{1} = 0$, let q^* be the smallest cut quotient over the cut quotients of all threshold cuts based on $\mathbf{x}$. Then*

$$q^* \leq \sqrt{2\Delta \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}}.$$

Proof: Assume w.l.o.g. that the vertices of the graph are numbered such that the entries of $\mathbf{x}$ occur in non-increasing order: for $i < j$, $x_i \geq x_j$. Let B be the generalized sum Laplacian as described above.

We start with two facts about quadratic terms of generalized Laplacians and sum Laplacians. In the expressions below, let $\mathbf{z}$ be any real vector. First, the following fact is well known:

$$\mathbf{z}^T A \mathbf{z} = \sum_{(i,j)\in E(G)} w_{ij}\, (z_i - z_j)^2 \tag{1}$$

Second,

$$
\begin{aligned}
\left(\mathbf{z}^T A \mathbf{z}\right)\left(\mathbf{z}^T B \mathbf{z}\right) &= \left(\sum_{(i,j)\in E(G)} w_{ij}\ (z_i - z_j)^2\right)\left(\sum_{(i,j)\in E(G)} w_{ij}\ (z_i + z_j)^2\right) \\
&= \left(\sum_{(i,j)\in E(G)} \left(\sqrt{w_{ij}}\ |z_i - z_j|\right)^2\right)\left(\sum_{(i,j)\in E(G)} \left(\sqrt{w_{ij}}\ |z_i + z_j|\right)^2\right) \\
&\geq \left(\sum_{(i,j)\in E(G)} w_{ij}\ |z_i^2 - z_j^2|\right)^2, \qquad (2)
\end{aligned}
$$

where the third line follows from the Cauchy-Schwarz inequality.

It is useful to give a high-level outline of the proof here before proceeding: we have just shown that the product $\left(\mathbf{x}^T A \mathbf{x}\right)\left(\mathbf{x}^T B \mathbf{x}\right)$ provides a connection between $\mathbf{x}^T A \mathbf{x}$ (which is expressed in terms of a weighted sum of squares of differences across edges) and a weighted sum of differences of the squares of the values at the ends of edges. The second sum telescopes, and can be neatly divided up in terms of subintervals of the the interval from x_i to x_j. This will allow us to break an edge up into a number of pieces corresponding to the number of thresholds (and hence cuts) that it crosses. We will rewrite the last sum in (2) as a weighted sum of cut quotients to prove the theorem. However, two issues must be addressed: First, the weighted sum will involve cut quotients, which use the size of the smaller shore of the cut as a denominator. Second, any edge that crosses zero is a potential problem for the application of telescoping. In the argument below, we break the contribution of an edge into (positive) contributions for subintervals. For an edge (i, j) crossing the zero point, the sum of the contributions could be bigger than the difference $w_{ij}\ |x_i^2 - x_j^2|$. This could violate the inequalities used to show the upper bound. Therefore it is useful to make two changes: We shift the values of $\mathbf{x}$ so that $x_{\lceil \frac{n}{2} \rceil} = 0$; and we modify G by breaking any edge that crosses the zero point into two parts, one part from x_i to a vertex with value zero, and one part from the zero vertex to x_j; each of these parts is assigned weight w_{ij}. The next section shows that these changes don't affect the preceding upper bound much.

Let G' be the graph modified as specified in the previous paragraph; G' has Laplacian A'. Let $\mathbf{z}$ be any nonzero vector such that $z_i \geq z_j$ for all $i < j$ and $z_{\lceil \frac{n}{2} \rceil} = 0$. Then with respect to equation (1), $\mathbf{z}^T A' \mathbf{z}$ and $\mathbf{z}^T A \mathbf{z}$ differ only in the terms for edges that go from some vertex $i < \lceil \frac{n}{2} \rceil$ to some vertex $j > \lceil \frac{n}{2} \rceil$. Note that for each such edge we have

$$(z_i - z_j)^2 = z_i^2 + z_j^2 - 2 z_i z_j > z_i^2 + z_j^2 = (z_i - 0)^2 + (0 - z_j)^2,$$

where the inequality holds because z_i and z_j have opposite signs by our restriction on the ordering of $\mathbf{z}$ (the edge weight has been factored out of each expression). Thus we have that

$$\frac{\mathbf{z}^T A'\mathbf{z}}{\mathbf{z}^T\mathbf{z}} \leq \frac{\mathbf{z}^T A\mathbf{z}}{\mathbf{z}^T\mathbf{z}} \tag{3}$$

for any such vector.

Now consider the shifted version of $\mathbf{x}$: Let $\mathbf{y} = \mathbf{x} + \alpha\vec{1}$ where $\alpha = -x_{\lceil \frac{n}{2} \rceil}$. We have the following:

$$\frac{\mathbf{y}^T A\mathbf{y}}{\mathbf{y}^T\mathbf{y}} = \frac{(\mathbf{x}+\alpha\vec{1})^T A(\mathbf{x}+\alpha\vec{1})}{(\mathbf{x}+\alpha\vec{1})^T(\mathbf{x}+\alpha\vec{1})} = \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}+\alpha^2 n} \leq \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}},$$

where the second equality follows from the restriction $\mathbf{x}^T\vec{1} = 0$ from the theorem statement, and from the fact that $\vec{1}$ is the (simple) zero eigenvalue for any (generalized) Laplacian. Since $\mathbf{y}$ meets the restrictions on $\mathbf{z}$ in the preceding paragraph, we can combine this result with inequality (3) to get

$$\mathbf{y}^T A'\mathbf{y} \leq \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \cdot \mathbf{y}^T\mathbf{y}. \tag{4}$$

We can perform a similar analysis for B', the sum Laplacian of G':

$$\frac{\mathbf{y}^T B'\mathbf{y}}{\mathbf{y}^T\mathbf{y}} \leq \frac{\mathbf{y}^T B\mathbf{y}}{\mathbf{y}^T\mathbf{y}} = \frac{\mathbf{y}^T(2D-A)\mathbf{y}}{\mathbf{y}^T\mathbf{y}} < \frac{\mathbf{y}^T(2D)\mathbf{y}}{\mathbf{y}^T\mathbf{y}} \leq \frac{\mathbf{y}^T(2\Delta I)\mathbf{y}}{\mathbf{y}^T\mathbf{y}} = 2\Delta.$$

The first inequality follows from the fact that A' is positive semidefinite, and that $\mathbf{y}$ is not a multiple of the "all ones" vector, the only zero eigenvalue of A'. The second inequality replaces the degree matrix with ΔI; this follows because the only vertex in G' that could have degree greater than Δ is $\lceil \frac{n}{2} \rceil$; however, the corresponding entry of $\mathbf{y}$ is 0 and the inequality holds. We thus have that

$$\mathbf{y}^T B'\mathbf{y} \leq 2\Delta \cdot \mathbf{y}^T\mathbf{y}. \tag{5}$$

Combining inequalities (2), (4), and (5), we get

$$2\Delta \cdot \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \cdot \left(\mathbf{y}^T\mathbf{y}\right)^2 \geq (\mathbf{y}^T B'\mathbf{y})\,(\mathbf{y}^T A'\mathbf{y}) \geq \left(\sum_{(i,j)\in E(G')} w_{ij}\,|y_i^2 - y_j^2|\right)^2.$$

Since only nonnegative values are involved, we can take the square root of the terms above. Further, since no edges cross the zero point, we can rewrite the summation to eliminate the absolute value signs. This gives the following:

$$\sqrt{2\Delta\frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}}} \cdot \left(\mathbf{y}^T\mathbf{y}\right) \geq \sum_{\substack{(i,j)\in E(G') \\ i<j\leq\lceil\frac{n}{2}\rceil}} w_{ij}(y_i^2 - y_j^2) + \sum_{\substack{(i,j)\in E(G') \\ \lceil\frac{n}{2}\rceil\leq i<j}} w_{ij}(y_j^2 - y_i^2). \tag{6}$$

The rest of the proof essentially follows Mohar's proof; the main distinction is that Mohar only worked with the positive side of the vector he considered. We include both sides of the vector.[1] We'll actually only show the proof for the positive part of the vector, however. The argument for the negative half is symmetric and left as an exercise.

We need some notation before we can finish the proof. Note that the y_i's may not be distinct. Assume that there are k distinct values in the subvector consisting of entries y_1 through $y_{\lceil \frac{n}{2} \rceil}$, and denote them as $t_1 > t_2 > \ldots > t_{k-1} > t_k = 0$. Let δV_i be the total weight of the edges (k, l) in G' such that $y_k \geq t_i$ and $y_l < t_i$; that is, δV_i is the weight of the edges crossing the cut at threshold t_i. Let $V_i = \{j \in V(G') \,|\, y_j \geq t_i\}$ (for simplicity of notation below, let $V_0 = \emptyset$). Finally, let q_i be the quotient cut that separates V_i from the rest of the graph, and let q^* be the minimum quotient cut produced by vector $\mathbf{y}$. The definition for cut quotient thus can be stated as follows:

$$q_i = \frac{\delta V_i}{|V_i|}. \tag{7}$$

Note that, by the construction of G' and $\mathbf{y}$, the values for the q_i's and q^* are unchanged if the definitions are applied to G and $\mathbf{x}$.

Consider the following calculation:

$$\begin{aligned}
\sum_{\substack{(i,j) \in E(G') \\ i < j \leq \lceil \frac{n}{2} \rceil}} w_{ij} \, (y_i^2 - y_j^2) &= \sum_{i=1}^{k-1} \delta V_i \, (t_i^2 - t_{i+1}^2) && (8) \\
&= \sum_{i=1}^{k-1} q_i \, |V_i| \, (t_i^2 - t_{i+1}^2) && (9) \\
&\geq q^* \sum_{i=1}^{k-1} |V_i| \, (t_i^2 - t_{i+1}^2) && (10) \\
&= q^* \sum_{i=1}^{k-1} (|V_i| - |V_{i-1}|) \, t_i^2 && (11) \\
&= q^* \sum_{i=1}^{\lceil \frac{n}{2} \rceil} y_i^2. && (12)
\end{aligned}$$

The first step in deriving equation (8) is the application of telescoping: Let $y_i = t_l$ and $y_j = t_m$. Then $y_i^2 - y_j^2 = \sum_{i=l}^{m-1} (t_i^2 - t_{i+1}^2)$. This sum is regrouped with respect to the differences $t_i^2 - t_{i+1}^2$; each such difference is weighted by a factor equal to the

[1] Note that when $x_{\lceil \frac{n}{2} \rceil}$ is the minimum or maximum value of $\mathbf{x}$, one of the sums on the right hand side of (6) will be zero.

weight of the edges crossing that threshold. Equality (9) follows by an application of (7). The inequality (10) then follows from the definition of $q*$. Equation (11) is a reordering of the preceding sum based on noting that t_i^2 occurs in (10) only in the expressions for $|V_i|$ and $|V_{i-1}|$; recall that $t_k = 0$. Finally, $|V_i| - |V_{i-1}|$ is the number of vertices with value t_i; equation (12) reintroduces the corresponding values from $\mathbf{y}$, including any zero values with indices less than or equal to $\lceil \frac{n}{2} \rceil$.

As noted before, the argument for the negative half of $\mathbf{y}$ is symmetric. Combining the two results (remember that $y_{\lceil \frac{n}{2} \rceil} = 0$ and that $\mathbf{y}^T\mathbf{y} = \sum_{i=1}^{n} y_i^2$) and applying (6) completes the proof.

□