November 7, 2005
Ross M. Richardson

Edge Isoperimetric Inequalities

1 Four Questions

Recall that in the last lecture we looked at the problem of isoperimetric inequalities in the hypercube, Q_n. Our notion of boundary was that of vertex boundary, defined by $\delta(G) = \{v \notin S \mid v \sim u, u \in S\}$. We found that when trying to minimize the vertex boundary while holding the size of our set S fixed, the (near) Hamming balls characterize those S with minimal vertex boundary.

In this lecture, we consider a different notion of boundary based on edges. We define the *edge boundary* of S to be $\partial(S) = \{\{u, v\} \mid u \in S, v \notin S\}$. This is exactly the set of edges required to disconnect S from any vertex not in S.

1.1 Four Problems

We shall look at four related isoperimetric problems which all utilize the concept of edge boundary.

1.1.1 Question 1

We begin with the simplest question: Given a fixed positive integer m, what is the smallest edge boundary for a set of m vertices? We may formalize this question by defining

$$g(m) = \min_{\substack{S \subseteq V(Q_n) \\ |S| = m}} |\partial(S)|.$$

We may then ask: Can we place bounds on $g(m)$? Can we characterize those subsets S which achieve the value of $g(m)$? Whereas in the case of vertex boundary, the optimal subsets S were found to be (near) Hamming balls, we shall see that for the edge isoperimetric problem these optimal sets are (near) subcubes. Because the problem of finding a set S with minimum edge boundary is dual to the problem of maximizing the number of edges between vertices in S, this result makes intuitive sense.

1.1.2 Question 2

For a subset $S \subseteq V(Q_n)$ of size m, let E_i be all those edges in $E(Q_n)$ whose vertices differ in the i^{th} coordinate. Set $\partial_i(S) = \partial(S) \cap E_i$. This provides a partition of $\partial(S)$, with

$$\sum_i |\partial_i(S)| = |\partial(S)|.$$

We may then define our function of interest as

$$g_1(m) = \min_{\substack{S \subseteq V(Q_n) \\ |S|=m}} \max_i |\partial_i(S)|.$$

Thus we have that for every subset S of size m, $\partial_i(S) \geq g_1(m)$ for some i. Trivially we obtain the relation $g_1(m) \geq g(m)/n$. Can we further characterize $g_1(m)$? This question was first proposed by Ben-Or and Linial [1]

1.1.3 Question 3

An alternate direction is to look at subgraphs induced by a subset S of vertices. For $S \subseteq V(Q_n)$, we denote by $Q_n[S]$ the subgraph of Q_n *induced* by S, that is, the graph on vertex set S and containing all edges of $E(Q_n)$ for which both endpoints are in S. For positive integer m, define

$$g_2(m) = \min_{\substack{S \subseteq V(Q_n) \\ |S|=m}} |E\left(Q_n[S]\right)|.$$

Note that if $m \leq 2^{n-1}$ we have $g_2(m) = 0$, since the hypercube is bipartite. However, when $m = 2^{n-1} + 1$, we obtain a sharp jump in the number of edges of $Q_n[S]$.

1.1.4 Question 4

Let $\Delta(G)$ denote the maximum degree of a graph G. We may again look at induced subgraphs of the hypercube, but measure their maximum degree. Formally we define

$$g_3(m) = \min_{\substack{S \subseteq V(Q_n) \\ |S|=m}} \Delta(Q_n[S]).$$

Once again, we must restrict our attention to $m > 2^{n-1}$. Can we characterize $g_3(m)$? Currently, both $g_2(m)$ and $g_3(m)$ are not fully understood even for $m = 2^{n-1} + 1$.

2 Theorems

We shall explore the first two questions.

2.0.5 On the First Question

Our first question asks simply what subsets have minimal edge boundaries. If we let m be a power of 2 then our answer is simple: a subcube of size m. If m is not an integer the answer is not too different, but unfortunately it becomes a bit less beautiful to both state and prove. Nonetheless, we have the following Theorem:

Theorem 1. *Fix a positive integer m. For $0 \leq \gamma \leq m$, we may express γ by its binary expansion $\gamma = \sum_{i=0}^{n-1} \gamma_i 2^i$. Let $S = \{(\gamma_0, \ldots, \gamma_{n-1}) \mid 0 \leq \gamma \leq m\}$ be the set of all binary expansions of positive integers $\gamma \leq m$. Then S achieves the minimum edge boundary of all subsets with m vertices.*

Proof. By Induction. This is left as **Exercise 1**. □

If we sacrifice an exact description, we obtain a much cleaner bound.

Theorem 2. *For $S \subseteq V(Q_n)$ with $|S| = m$,*

$$|\partial(S)| \geq m(n - \log_2 m).$$

In order to prove Theorem 2, we shall require an additional result. Recall that the average degree $\bar{d}$ of a graph G is defined as $\bar{d} = \frac{\sum_{v \in V(G)} d_G(v)}{|G|}$.

Theorem 3. *Let G be a subgraph of $Q_n[S]$ with average degree $\bar{d}$. Then*

$$|V(G)| \geq 2^{\bar{d}}.$$

Let us first show how Theorem 3 implies Theorem 2. Note that $|\partial(S)| = m(n - \bar{d})$, by definition. By Theorem 3, we have $2^{\bar{d}} \leq m$ or $\bar{d} \leq \log_2 m$. Combining these two observations yields

$$|\partial(S)| \geq m(n - \log_2 m).$$

We thus need only prove Theorem 3.

Proof (Theorem 3). We prove this by induction. We view Q_n as composed of two $(n-1)$cubes, which we label $Q^{(1)}$ and $Q^{(2)}$. We let G_1 and G_2 be the intersection of G with these two subcubes, and set $m_i = |V(G_i)|$. Without loss of generality, assume $0 \leq m_1 \leq m_2$. Finally assume there are s edges of G between G_1 and G_2. Note that for each vertex in G_1 there can be at most one edge adjacent to this vertex crossing to G_2, so $s \leq m_1$.

We may conclude by induction that

$$m_i \log_2 m_i \geq \sum_{v \in V(G_i)} d_{G_i}(v) = \sum_{v \in V(G_i)} d_G(v) - s, \qquad i = 1, 2.$$

Observe that

$$m_1 \log_2 m_1 + m_2 \log_2 m_1 + 2s \geq \sum_{v \in V(G)} d_G(v).$$

Thus we have, noting that $m_2 \geq m_1$

$$\begin{aligned} m \log_2 m = (m_1 + m_2)\log_2(m_1 + m_2) &\geq m_1 \log_2 m_1 + m_2 \log_2 m_2 + 2m_1 \\ &\geq \sum_{v \in V(G)} d_G(v). \end{aligned}$$

Here, the first inequality is a fact which is proved at the end of these notes (where one can see why the base two logarithm is required), and the second inequality follows simply from the calculations done above and observing that $m_1 \geq s$. Thus, the proof concludes. □

2.1 On the second question

To begin to approach the second question, we shall introduce the notion of *Boolean functions*. A Boolean function is simply a function on $0/1$ strings of length n which is either zero or one. It shall be of some utility to view each coordinate of the length n string as a separate $0/1$ variables, which we shall denote $x_1, \ldots, x_n$. Let us now look at three important examples.

$$\begin{aligned} &\text{Parity:} && f_1(x_1, \ldots, x_n) = \sum_i x_i \pmod 2 \\ &\text{Projection:} && f_2(x_1, \ldots, x_n) = x_1 \\ &\text{Majority:} && f_3(x_1, \ldots, x_n) = \begin{cases} 1 & \sum_i x_i \geq n/2, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

To analyze these examples we shall associate to each Boolean function a game. Each game has n players associated to one each of the variables $x_1, \ldots, x_n$. Each player flips a fair coin to decide the value of x_i. The value of the game is just the value of $f(x_1, \ldots, x_n)$. But what if one of the players acts intelligently, choosing his value deterministically in order to influence the value of the game? We can then ask, how much influence can a single player have on the value of the game?

To formalize this question, let us fix some index i, and form the Boolean function $f_i(0)$ on the variables $x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ by setting

$$f_i(0)(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = f(x_1, x_2, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n).$$

We form $f_i(1)$ similarly. We may then define the *influence* $I_f(i)$ of the i^{th} player to be

$$I_f(i) = P(f_i(0) \neq f_i(1)).$$

Observe that the parity function satisfies $I_f(i) = 1$, because flipping one bit in a string changes the parity of the entire string. For the projection function, observe that $I_f(i)$ is 1 if $i = 1$ and 0 otherwise, i.e. the first player completely determines the value of f. The majority function is more complicated, and it is a nontrivial fact that the influence of the i^{th} player is $1/\sqrt{n}$.

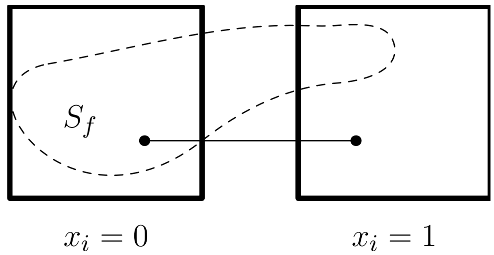

Figure 1: An i^{th} coordinate boundary edge corresponds to a place of influence for player i.

Say that the coin-flipping game is *fair* if the probability that $f = 0$ is $1/2$. We can show that for any fair game, there must exist some player with influence at least $1/n$. To see this, we need to relate our game to the hypercube.

Observe that a Boolean function f determines a partition of Q_n. We denote by S_f the set of all vertices of the hypercube for which $f = 1$. If $P(f = 0) = 1/2$ then S_f consists of exactly half of the vertices of Q_n. We again note that for a fixed i, we can view the Q_n as two copies of Q_{n-1} (denoted by $Q^{(1)}$ and $Q^{(2)}$) joined by edges along the i^{th} coordinate. Observe that the number of strings for which changing i's value changes the value of f is simply the number edges from $\partial(S_f)$ which cross from $Q^{(1)}$ to $Q^{(2)}$, denoted by $\partial_i(S)$ (see figure 1). Thus, the influence of player i is given by

$$I_f(i) = \frac{|\partial_i(S_f)|}{2^{n-1}}.$$

Summing over all i, we have

$$\sum_i I_f(i) = \frac{|\partial(S_f)|}{2^{n-1}}.$$

Finally, we may apply Theorem 2 with $m = |S_f| = 2^{n-1}$ to give a lower bound on $|\partial(S_f)|$, which forces the right hand side to be at least one. Thus, $I_f(i) \geq 1/n$ for some i.

Indeed, more is true. Kahn, Kalai, and Linial [3] have shown that for any fair game there exists some player i such that

$$I_f(i) \geq c\frac{\log n}{n},$$

where here c is some fixed constant.

3 Appendix

Lemma 1. *If $0 \leq x \leq y$ then*

$$(x + y)\log_2(x + y) \geq x \log_2 x + y \log_2 y + 2x.$$

Proof. Assume $x > 0$, otherwise the result is trivial. Let $\gamma = y/x$. Then

$$\begin{aligned}(x+y)\log_2(x+y) &\geq x\log_2 x(1+\gamma) + y\log_2 y(1+1/\gamma)\\ &= x\log_2 x + y\log_2 y + x\log_2(1+\gamma) + y\log_2(1+1/\gamma)\\ &= x\log_2 x + y\log_2 y + x(\log_2(1+\gamma) + \gamma\log_2(1+1/\gamma))\end{aligned}$$

To see that $\log_2(1+\gamma) + \gamma\log_2(1+1/\gamma) \geq 2$ we observe that

$$\begin{aligned}\log_2(1+\gamma) + \gamma\log_2(1+1/\gamma) &\geq \log_2(1+\gamma) + \log_2(1+1/\gamma)\\ &\geq \log_2(1+\gamma)(1+1/\gamma)\\ &\geq \log_2(2+\gamma+1/\gamma) = 2.\end{aligned}$$

Where here we use the basic fact that $\gamma + 1/\gamma \geq 2$. $\square$

References

[1] Ben-Or, M. and N. Linial. "Collective Coin Flipping, Robust Voting Schemes and Minima of Banzhaf Values". FOCS 1985: 408-416.

[2] Chung, Fan. "Edge Isoperimetric Inequalities." Draft, 2005.

[3] Kahn, J., Kalai, G. and N. Linial. "The Influence of Variables on Boolean Functions". FOCS 1988:68-80.