Math 262B Lecture Notes Week 3

Prof. Fan Chung Graham

Compiled by Jia Mao

1 Addressing Problem on Digraphs

Let $G(V, E)$ be a digraph with vertex set $V = \{v_1, v_2, ..., v_n\}$ and edge set E. We assume that G is *strongly connected*, meaning there is a directed path between any pair of vertices (v_i, v_j). For $v_i, v_j \in V$, the *distance* $d(v_i, v_j) = d_{ij}$ is defined to be the length of a shortest path from v_i to v_j. Let $D = \{d_{ij}\}$ be the $n \times n$ *distance matrix* of G.

Notice that in general d_{ij} is not symmetric and also that we will not be able to use *trees* to help our analysis. This indicates the relative difficulty of the directed case vs. the undirected case.

Again we consider an address assignment to the vertices $f : V \to \{0, 1, *\}^k$. For any $f(v_i) = a = (a_1, ..., a_k)$ and $f(v_j) = b = (b_1, ..., b_k)$, define the *hamming distance* $d_h(a, b) = |\{i : a_i = 0, b_i = 1\}|$. An address assignment is called *good* if it is *distance preserving*, i.e. $d_G(v_i, v_j) = d_h(a, b)$ for $\forall f(v_i) = a, f(v_j) = b$. Our goal is to find a good address assignment with k minimized. We denote $q(G)$ to be the minimum k for which we have a good address assignment for G.

Theorem 1.1 *For any digraph G, $q(G) \leq n(n-1)$.*

Proof of Theorem 1.1: Notice that an address assignment f can also be described by a $n \times k$ matrix A_f where row i represents the address of v_i. To construct a good address assignment A_f with $k = n(n-1)$, split A into n blocks of $(n-1)$ columns each, with each block handling the i^{th} row of the distance matrix D. Our assumption that G is strongly connected guarantees that every entry d_{ij} in D is a non-negative finite integer.

For the i^{th} block, we view it as a matrix of dimension $n \times (n-1)$, set the entries as

$$a_{js} = \begin{cases} 0 & \text{if } j = i \\ 1 & \text{if } j \neq i \text{ and } j-1 \leq s \leq d_{i,j} + j - 2 \\ * & \text{otherwise} \end{cases}$$

This gives us a good assignment with $k = n(n-1)$ address length.□

Theorem 1.2 *For directed cycles C_n, we have $q(C_n) > \mu n^{3/2}$ for some constanct $\mu > 0$.*

Proof of Theorem 1.2: Exercise.

2 Hypercubes

First we recall the definition of *hypercubes* from Lecture note week 1:

Definition 2.1 *Q_n(n-cube) is a graph with $V(Q_n) = \{0,1\}^n$, i.e. $V(Q_n) = \{a : a$ is a binary string of length $n\}$, and $a \sim b$ if and only if they differ in exactly one coordinate.*

Definition 2.2 *Hamming distance between u and v, $d_{Q_n}(u,v)$, is the number of coordinates in which u and v differ in Q_n.*

So far we have been looking at the following two problems:

Question 1: Given graph G, does G have an *embedding* into a hypercube Q_n that is distance-preserving?

Question 2: If affirmative, what is the minimum dimension of the hypercube? (ref **Karzanov 85**)

For question 1, we do know two necessary conditions for the given graph G to be embeddable into a hypercube.

1. G has to be bipartite.
2. For any edge $\{u,v\}$ and any vertex w, it should hold that $d(w,v) = d(w,u) + d(u,v)$ or $d(w,u) = d(w,v) + d(u,v)$.

3 Graph Embedding

One of the classical problems on graph embedding is when the host graph H is a path. Namely, given $G = (V, E)$, we consider the embeddings $f : V \to \{1, 2, ..., n\} = [1, n]$. There are a few related quantities we would like to compute for any embedding given $G = (V, E)$:

Definition 3.1 *For a particular f, the bandwidth is defined to be the* maximum stretch

$$B(f) = \max_{x \sim y} |f(x) - f(y)|$$

Bandwidth (a.k.a. dilation)for the graph G is defined as

$$B(G) = \min_{f} B(f)$$

In other words, the *bandwidth* problem is to arrange the vertices of a graph into a line (of integers) such that the maximum stretch of edges of the graph is minimized.

Definition 3.2 *For a particular f, the cutwidth is defined to be*

$$CW(f) = \max_{i} |\{u, v\} \in E : f(u) \leq i < f(v)|$$

Bandwidth (a.k.a. congestion) for the graph G is defined as

$$CW(G) = \min_{f} CW(f)$$

In other words, the *cutwidth* problem is to arrange vertices in a line so that the maximum number of edges crossing the i_th place, over all i, is minimized.

Definition 3.3 *Linear arrangement for the graph G is defined as*

$$Sum(G) = \min_{f} \sum_{x \sim y} |f(x) - f(y)|$$

Compared with bandwidth and cutwidth which concern about the *extreme* case, the linear arrangement concerns about the *average* case.

We may also extend the embedding to the case when the host graph H is no longer discrete, namely the real line $H = \mathcal{R}$, and we consider $f : V \to \mathcal{R}$. It is of interest to look at $\Sigma_{x \sim y}(f(x) - f(y))^2$ in both worlds in order to get a feel of the bandwidth and cutwidth problems. This quantity, however, has interesting connection with the *combinatorial laplacian* of the original graph G.

4 Combinatorial Laplacian

There are a number of different views we can take to define the *combinatorial laplacian* of a given graph G. Different choice of them may be helpful in different contexts.

View 1:
Let $A_{n\times n}$ be the adjacency matrix of G, $D_{n\times n}$ be the diagonal degree matrix (i.e. degree of each vertex is on the diagonal), then the combinatorial laplacian of G is defined as $L = D - A$.

Notice that in $L = (l_{ij})$,

$$l_{ij} = \begin{cases} deg(i) & \text{for } i = j \\ -1 & \text{for } i \neq j \text{ and } i \sim j \\ 0 & \text{otherwise} \end{cases}$$

Lemma 4.1

$$\sum_{x\sim y}(f(x) - f(y))^2 = \langle f, Lf\rangle$$

Proof of Lemma 4.1: Let us first examine Lf, notice the element corresponding to vertex v

$$Lf(v) = deg(v)f(v) - \sum_{u\sim v} f(u) = \sum_{u\sim v}(f(v) - f(u))$$

Then we have

$$\begin{aligned} \langle f, Lf\rangle &= \textstyle\sum_v f(v)Lf(v) \\ &= \textstyle\sum_v f(v)\sum_{u\sim v}(f(v) - f(u)) \\ &= \textstyle\sum_{u\sim v}[f^2(v) - 2f(u)f(v) + f^2(u)] \\ &= \textstyle\sum_{u\sim v}[f(u) - f(v)]^2 \end{aligned}$$

This proves Lemma 4.1.□

View 2:
Let $B_{n\times m}$ be the incidence matrix of G, then the combinatorial laplacian of G is defined as $L = BB^T$.
In B, rows are indexed by V and columns indexed by E. For each edge $e = u, v \in E$ corresponding to a column, the entry where the row that corresponds to u is 1, and the entry where the row that corresponds to v is -1.

For undirected graph G, the 1 and -1 is assigned arbitrarily.

View 3:
Homology theory.

View 4:
Differential geometry.

One important feature of L is that because it's symmetric, it can be diagonalized. We will discuss how we can exploit this nice property in the next few lectures.