CS267: Notes for Lecture 23(a), Apr 9, 1996

Spectral partitioning - Statement and Proof of Theorem 1

Theorem 1. Given a graph G, its associated matrices In(G) and L(G) have the following properties.

1. L(G) is a symmetric matrix. This means the eigenvalues of L(G) are real, and its eigenvectors are real and orthogonal.
2. Let e=[1,...,1]', where ' means transpose, i.e. the column vector of all ones. Then L(G)*e = 0.
3. In(G)*(In(G))' = L(G). This is independent of the signs chosen in each column of In(G).
4. Suppose L(G)*v = lambda*v, where v is nonzero. Then

$$\text{lambda} = \text{norm}(\text{In(G)'*v})^2 \ / \ \text{norm}(v)^2 \quad \text{where } \text{norm}(z)^2 = \text{sum}_i \ z(i)^2$$
$$= \text{sum}_{\{\text{all edges } e=(i,j)\}} \ (v(i)-v(j))^2 \ / \ \text{sum}_i \ v(i)^2$$

5. The eigenvalues of L(G) are nonnegative: 0 <= lambda_1 <= lambda_2 <= ... <= lambda_n.
6. The number of of connected components of G is equal to the number of lambda_i equal to 0. In particular, lambda_2 != 0 if and only if G is connected.

Proof of part 1. Symmetry follows from the definition of L(G): Since G is an undirected graph, (i,j) is an edge if and only if (j,i) is an edge.

Proof of part 2. The i-th entry of L(G)*e is just the sum of the entries of the i-th row of L(G). This equals the degree of node i, L(G)(i,i), minus 1 for each incident edge (L(G)(i,j), or exactly zero.

Proof of part 3.

$$(\text{In(G)*(In(G))'})(i,i) = \text{sum}_{\{\text{all edges } e, \text{ such that } i \text{ is an endpoint of } e\}} \ (+-1)^2 = \text{degree of node } i$$

and

$$(\text{In(G)*(In(G))'})(i,j) = \text{sum}_{\{\text{all edges } e = (i,j)\}} \ (-1)*(+1) = -1 \text{ if an edge } e=(i,j) \text{ exists}$$

Proof of part 4. Suppose L(G)*v = lambda*v, where lambda is an eigenvalues and v is a nonzero eigenvector. Then v'*L(G)*v = lambda*v'*v, where v'*v is a positive scalar. Thus

$$\begin{aligned}\text{lambda} &= (\ v'*\ L(G)\ *v\ )\ /\ (\ v'*v\ ) \\ &= (\ v'*\ (In(G)*(In(G))')\ *v\ )\ /\ (\ v'*v\ ) \\ &= (\ v'*\ In(G))\ *\ (In(G))'\ *v\ )\ /\ (\ v'*v\ ) \\ &= (\ y'\ )\ *\ (\ y\ )\ /\ (\ v'*v\ ) \quad \text{where } y = In(G)'*v \\ &= \text{sum}_e\ y(e)^2\ /\ \text{sum}_i\ v(i)^2\end{aligned}$$

If edge e = (i,j), it is easy to see by construction that y(e) = v(i)-v(j) or its negative, depending on the arbitrary choice of signs in column e of In(G). Thus $y(e)^2 = (v(i)-v(j))^2$, independent of the choice of sign.

Proof of part 5. By part 4, each eigenvalue lambda is the quotient of two nonnegative quantities, and so must be nonnegative.

Proof of part 6. For lambda to equal 0, each y(e) in the expression

$$\text{lambda} = \text{sum}_e\ y(e)^2 \ /\ \text{sum}_i\ v(i)^2$$

must be zero. This means v(i)=v(j) for each edge e=(i,j). Starting with any node i and applying the fact v(i)=v(j) repeatedly, one can see that any node k reachable from i also satisfies v(k)=v(i)=c. In other words, the eigenvector v is a constant c on each connected component. Since L(G) is symmetric, the number of independent eigenvectors corresponding to lambda=0 is equal to the number of eigenvalues equal to 0. If there are exactly d connected components, there are exactly d independent eigenvectors, since choosing d constants c(1),...,c(d) (one for each connected component) determines each eigenvector uniquely.