Chapter 16

Spectral Graph Theory

Daniel Spielman

Yale University

16.1 Introduction

Spectral graph theory is the study and exploration of graphs through the eigenvalues and eigenvectors of matrices naturally associated with those graphs. It is intuitively related to attempts to understand graphs through the simulation of processes on graphs and through the consideration of physical systems related to graphs. Spectral graph theory provides many useful algorithms, as well as some that can be rigorously analyzed. We begin this chapter by providing intuition as to why interesting properties of graphs should be revealed by these eigenvalues and eigenvectors. We then survey a few applications of spectral graph theory.

The figures in this chapter are accompanied by the Matlab code used to generate them.

16.2 Preliminaries

We ordinarily view an undirected graph[1] G as a pair (V, E), where V denotes its set of vertices and E denotes its set of edges. Each edge in E is an unordered pair of vertices, with the edge connecting distinct vertices a and b written as (a, b). A weighted graph is a graph in which a weight (typically a real number) has been assigned to every edge. We denote a weighted graph by a triple (V, E, w), where (V, E) is the associated unweighted graph, and w is a function from E to the real numbers. We restrict our attention to weight functions w that are strictly positive. We reserve the letter n for the number of vertices in a graph. The *degree* of a vertex in an unweighted graph is the number of edges in which it is involved. We say that a graph is *regular* if every vertex has the same degree, and d-regular if that degree is d.

We denote vectors by bold letters, and denote the ith component of a vector $\boldsymbol{x}$ by $\boldsymbol{x}(i)$. Similarly, we denote the entry in the ith row and jth column of a matrix M by $M(i, j)$.

If we are going to discuss the eigenvectors and eigenvalues of a matrix M, we should be sure that they exist. When considering undirected graphs, most of the matrices we consider are symmetric, and thus they have an orthonormal basis of eigenvectors and n eigenvalues, counted with multiplicity. The other matrices we associate with undirected graphs are *similar* to symmetric matrices, and thus also have n eigenvalues, counted by multiplicity, and possess a basis of eigenvectors. In particular, these matrices are of the form MD^{-1}, where M is symmetric and D is a non-singular diagonal matrix. In this case, $D^{-1/2}MD^{-1/2}$ is symmetric, and we have

$$D^{-1/2}MD^{-1/2}\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i \quad \implies \quad MD^{-1}(D^{1/2}\boldsymbol{v}_i) = \lambda_i \left(D^{1/2}\boldsymbol{v}_i \right).$$

So, if $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ form an orthonormal basis of eigenvectors of $D^{-1/2}MD^{-1/2}$, then we obtain a basis (not necessarily orthonormal) of eigenvectors of MD^{-1} by multiplying these vectors by $D^{1/2}$. Moreover, these matrices have the same eigenvalues.

The matrices we associate with directed graphs will not necessarily be diagonalizable.

---

[1]Strictly speaking, we are considering simple graphs. These are the graphs in which all edges go between distinct vertices and in which there can be at most one edge between a given pair of vertices. Graphs that have multiple-edges or self-loops are often called multi-graphs.

16.3   The matrices associated with a graph

Many different matrices arise in the field of Spectral Graph Theory. In this section we introduce the most prominent.

16.3.1   Operators on the vertices

Eigenvalues and eigenvectors are used to understand what happens when one repeatedly applies an operator to a vector. If A is an n-by-n matrix having a basis of right-eigenvectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ with

$$A\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i,$$

then we can use these eigenvectors to understand the impact of multiplying a vector $\boldsymbol{x}$ by A. We first express $\boldsymbol{x}$ in the eigenbasis

$$\boldsymbol{x} = \sum_i c_i \boldsymbol{v}_i$$

and then compute

$$A^k \boldsymbol{x} = \sum_i c_i A^k \boldsymbol{v}_i = \sum_i c_i \lambda_i^k \boldsymbol{v}_i.$$

If we have an operator that is naturally associated with a graph G, then properties of this operator, and therefore of the graph, will be revealed by its eigenvalues and eigenvectors. The first operator one typically associates with a graph G is its *adjacency operator*, realized by its *adjacency matrix* A_G and defined by

$$A_G(i, j) = \begin{cases} 1 & \text{if } (i, j) \in E \\ 0 & \text{otherwise.} \end{cases}$$

To understand spectral graph theory, one must view vectors $\boldsymbol{x} \in \mathbb{R}^n$ as functions from the vertices to the Reals. That is, they should be understood as vectors in $\mathbb{R}^V$. When we apply the adjacency operator to such a function, the resulting value at a vertex a is the sum of the values of the function $\boldsymbol{x}$ over all neighbors b of a:

$$\left(A_G \boldsymbol{x}\right)(a) = \sum_{b:(a,b) \in E} \boldsymbol{x}(b).$$

This is very close to one of the most natural operators on a graph: the *diffusion operator*. Intuitively, the diffusion operator represents a process in which "stuff" or "mass" moves from vertices to their neighbors. As mass should be conserved, the mass at a given vertex is distributed evenly among

its neighbors. Formally, we define the *degree* of a vertex a to be the number of edges in which it participates. We naturally encode this in a vector, labeled $\boldsymbol{d}$:

$$\boldsymbol{d}(a) = |\{b : (a, b) \in E\}| ,$$

where we write $|S|$ to indicate the number of elements in a set S. We then define the *degree matrix* D_G by

$$D_G(a, b) = \begin{cases} \boldsymbol{d}(a) & \text{if } a = b \\ 0 & \text{otherwise.} \end{cases}$$

The *diffusion matrix* of G, also called the *walk matrix* of G, is then given by

$$W_G \overset{\text{def}}{=} A_G D_G^{-1}. \tag{16.1}$$

It acts on a vector $\boldsymbol{x}$ by

$$\left(W_G \boldsymbol{x}\right)(a) = \sum_{b:(a,b)\in E} \boldsymbol{x}(b)/\boldsymbol{d}(b).$$

This matrix is called the *walk matrix* of G because it encodes the dynamics of a random walk on G. Recall that a random walk is a process that begins at some vertex, then moves to a random neighbor of that vertex, and then a random neighbor of that vertex, and so on. The walk matrix is used to study the evolution of the probability distribution of a random walk. If $\boldsymbol{p} \in \mathbb{R}^n$ is a probability distribution on the vertices, then $W_G \boldsymbol{p}$ is the probability distribution obtained by selecting a vertex according to $\boldsymbol{p}$, and then selecting a random neighbor of that vertex. As the eigenvalues and eigenvectors of W_G provide information about the behavior of a random walk on G, they also provide information about the graph.

Of course, adjacency and walk matrices can also be defined for weighted graphs $G = (V, E, w)$. For a weighted graph G, we define

$$A_G(a, b) = \begin{cases} w(a, b) & \text{if } (a, b) \in E \\ 0 & \text{otherwise.} \end{cases}$$

When dealing with weighted graphs, we distinguish between the *weighted degree* of a vertex, which is defined to be the sum of the weights of its attached edges, and the *combinatorial degree* of a vertex, which is the number of such edges. We reserve the vector $\boldsymbol{d}$ for the weighted degree, so

$$\boldsymbol{d}(a) = \sum_{b:(a,b)\in E} w(a, b).$$

The random walk on a weighted graph moves from a vertex a to a neighbor b with probability proportional to $w(a, b)$, so we still define its walk matrix by equation (16.1).

16.3.2 The Laplacian Quadratic Form

Matrices and spectral theory also arise in the study of quadratic forms. The most natural quadratic form to associate with a graph is the *Laplacian*, which is given by

$$\boldsymbol{x}^T L_G \boldsymbol{x} = \sum_{(a,b) \in E} w(a,b)(\boldsymbol{x}(a) - \boldsymbol{x}(b))^2. \tag{16.2}$$

This form measures the smoothness of the function $\boldsymbol{x}$. It will be small if the function $\boldsymbol{x}$ does not jump too much over any edge. The matrix defining this form is the *Laplacian matrix* of the graph G,

$$L_G \stackrel{\text{def}}{=} D_G - A_G.$$

The Laplacian matrices of weighted graphs arise in many applications. For example, they appear when applying the certain discretization schemes to solve Laplace's equation with Neumann boundary conditions. They also arise when modeling networks of springs or resistors. As resistor networks provide a very useful physical model for graphs, we explain the analogy in more detail. We associate an edge of weight w with a resistor of resistance $1/w$, since higher weight corresponds to higher connectivity which corresponds to less resistance.

When we inject and withdraw current from a network of resistors, we let $\boldsymbol{i}_{ext}(a)$ denote the amount of current we inject into node a. If this quantity is negative then we are removing current. As electrical flow is a potential flow, there is a vector $\boldsymbol{v} \in \mathbb{R}^V$ so that the amount of current that flows across edge (a,b) is

$$\boldsymbol{i}(a,b) = \left(\boldsymbol{v}(a) - \boldsymbol{v}(b)\right)/r(a,b),$$

where $r(a,b)$ is the resistance of edge (a,b). The Laplacian matrix provides a system of linear equations that may be used to solve for $\boldsymbol{v}$ when given $\boldsymbol{i}_{ext}$:

$$\boldsymbol{i}_{ext} = L_G \boldsymbol{v}. \tag{16.3}$$

We refer the reader to [1] or [2] for more information about the connections between resistor networks and graphs.

16.3.3 The Normalized Laplacian

When studying random walks on a graph, it often proves useful to normalize the Laplacian by its degrees. The *normalized Laplacian* of G is defined by

$$N_G = D_G^{-1/2} L_G D_G^{-1/2} = I - D_G^{-1/2} A_G D_G^{-1/2}.$$

It should be clear that normalized Laplacian is closely related to the walk matrix of a graph. Chung's monograph on spectral graph theory focuses on the normalized Laplacian [3].

16.3.4   Naming the Eigenvalues

When the graph G is understood, we will always let

$$\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$$

denote the eigenvalues of the adjacency matrix. We order the eigenvalues of the Laplacian in the other direction:

$$0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n.$$

We will always let

$$0 = \nu_1 \leq \nu_2 \leq \cdots \leq \nu_n$$

denote the eigenvalues of the normalized Laplacian. Even though ω is not a Greek variant of w, we use

$$1 = \omega_1 \geq \omega_2 \geq \cdots \geq \omega_n$$

to denote the eigenvalues of the walk matrix. It is easy to show that $\omega_i = 1 - \nu_i$.

For graphs in which every vertex has the same weighted degree the degree matrix is a multiple of the identity; so, A_G and L_G have the same eigenvectors. For graphs that are not regular, the eigenvectors of A_G and L_G can behave very differently.

16.4   Some Examples

The most striking demonstration of the descriptive power of the eigenvectors of a graph comes from Hall's spectral approach to graph drawing [4]. To begin a demonstration of Hall's method, we generate the Delaunay graph of 200 randomly chosen points in the unit square.

```
xy = rand(200,2);
tri = delaunay(xy(:,1),xy(:,2));
elem = ones(3)-eye(3);
for i = 1:length(tri),
  A(tri(i,:),tri(i,:)) = elem;
end
A = double(A > 0);
gplot(A,xy)
```
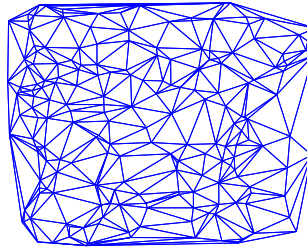
We will now discard the information we had about the coordinates of the vertices, and draw a picture of the graph using only the eigenvectors of its Laplacian matrix. We first compute the adjacency matrix A, the degree matrix D, and the Laplacian matrix L of the graph. We then compute the

eigenvectors of the second and third smallest eigenvalues of L, $\boldsymbol{v}_2$ and $\boldsymbol{v}_3$. We then draw the same graph, using $\boldsymbol{v}_2$ and $\boldsymbol{v}_3$ to provide the coordinates of vertices. That is, we locate vertex a at position $(\boldsymbol{v}_2(a), \boldsymbol{v}_3(a))$, and draw the edges as straight lines between the vertices.

```
D = diag(sum(A));
L = D - A;
[v,e] = eigs(L, 3, 'sm');
gplot(A,v(:,[2 1]))
```
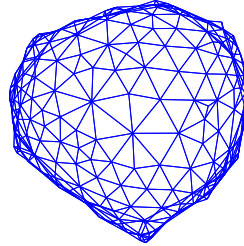
Amazingly, this process produces a very nice picture of the graph, in spite of the fact that the coordinates of the vertices were generated solely from the combinatorial structure of the graph. Note that the interior is almost planar. We could have obtained a similar, and possibly better, picture from the left-eigenvectors of the walk matrix of the graph.

```
W = A * inv(D);
[v,e] = eigs(W', 3);
gplot(A,v(:,[2 3]));
```
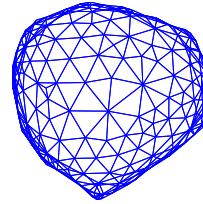
We defer the motivation for Hall's graph drawing technique to Section 16.7, so that we may first explore other examples.

One of the simplest graphs is the path graph. In the following figure, we plot the 2nd, 3rd, 4th, and 12th eigenvectors of the Laplacian of the path graph on 12 vertices. In each plot, the x-axis is the number of the vertex, and the y-axis is the value of the eigenvector at that vertex. We do not bother to plot the 1st eigenvector, as it is a constant vector.

```
A = diag(ones(1,11),1);
A = A + A';
D = diag(sum(A));
L = D - A;
[v,e] = eig(L);
plot(v(:,2),'o'); hold on;
plot(v(:,2));
plot(v(:,3),'o'); hold on;
plot(v(:,3));
. . .
```
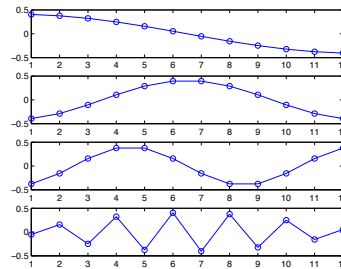
Observe that the 2nd eigenvector is monotonic along the path, that the second changes sign twice, and that the 12th alternates negative and positive. This can be explained by viewing these eigenvectors as the fundamental modes

of vibration of a discretization of a string. We recommend [5] for a formal treatment.

By now, the reader should not be surprised to see that ring graphs have the obvious spectral drawings. In this case, we obtain the ring from the path by adding an edge between vertex 1 and 12.

```
A(1,12) = 1; A(12,1) = 1;
D = diag(sum(A));
L = D - A;
[v,e] = eig(L);
gplot(A,v(:,[2 3]))
hold on
gplot(A,v(:,[2 3]),'o')
```
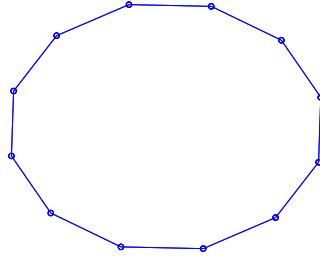
Our last example comes from the skeleton of the "Buckyball". This is the same as the graph between the corners of the Buckminster Fuller geodesic dome and of the seams on a standard Soccer ball.

```
A = full(bucky);
D = diag(sum(A));
L = D - A;
[v,e] = eig(L);
gplot(A,v(:,[2 3]))
hold on;
gplot(A,v(:,[2 3]),'o')
```
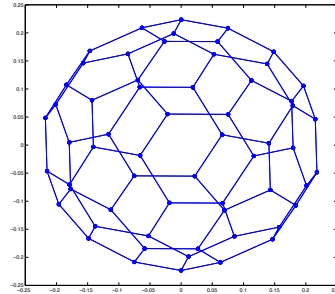
Note that the picture looks like a squashed Buckyball. The reason is that there is no canonical way to choose the eigenvectors v_2 and v_3. The smallest non-zero eigenvalue of the Laplacian has multiplicity three. This graph should really be drawn in three dimensions, using any set of orthonormal vectors v_2, v_3, v_4 of the smallest non-zero eigenvalue of the Laplacian. As this picture hopefully shows, we obtain the standard embedding of the Buckyball in $\mathbb{R}^3$.

```
[x,y] = gplot(A,v(:,[2 3]));
[x,z] = gplot(A,v(:,[2 4]));
plot3(x,y,z)
```
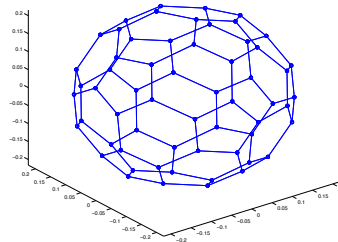
The Platonic solids and all vertex-transitive convex polytopes in $\mathbb{R}^d$ display similar behavior. We refer the reader interested in learning more about this phenomenon to either Godsil's book [6] or to [7].

16.5   The Role of the Courant-Fischer Theorem

Recall that the Rayleigh quotient of a non-zero vector $\boldsymbol{x}$ with respect to a symmetric matrix A is

$$\frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}.$$

The Courant-Fischer characterization of the eigenvalues of a symmetric matrix A in terms of the maximizers and minimizers of the Rayleigh quotient (see [8]) plays a fundamental role in spectral graph theory.

Theorem 3 (Courant-Fischer) *Let A be a symmetric matrix with eigenvalues $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$. Then,*

$$\alpha_k = \max_{\substack{S \subseteq \mathbb{R}^n \\ \dim(S)=k}} \min_{\substack{\boldsymbol{x} \in S \\ \boldsymbol{x} \neq 0}} \frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} = \min_{\substack{T \subseteq \mathbb{R}^n \\ \dim(T)=n-k+1}} \max_{\substack{\boldsymbol{x} \in T \\ \boldsymbol{x} \neq 0}} \frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}.$$

The maximum in the first expression is taken over all subspaces of dimension k, and the minimum in the second is over all subspaces of dimension $n-k+1$.

Henceforth, whenever we minimize of maximize Rayleigh quotients we will only consider non-zero vectors, and thus will drop the quantifier "$\boldsymbol{x} \neq 0$".

For example, the Courant-Fischer Theorem tells us that

$$\alpha_1 = \max_{\boldsymbol{x} \in \mathbb{R}^n} \frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \quad \text{and} \quad \alpha_n = \min_{\boldsymbol{x} \in \mathbb{R}^n} \frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}.$$

We recall that a symmetric matrix A is *positive semidefinite*, written $A \succeq 0$, if all of its eigenvalues are non-negative. From (16.2) we see that the Laplacian is positive semidefinite. Adjacency matrices and walk matrices of non-empty graphs are not positive semidefinite as the sum of their eigenvalues equals their trace, which is 0. For this reason, one often considers the *lazy random walk* on a graph instead of the ordinary random walk. This walk stays put at each step with probability $1/2$. This means that the corresponding matrix is $(1/2)I + (1/2)W_G$, which can be shown to positive semidefinite.

16.5.1   Low-Rank Approximations

One explanation for the utility of the eigenvectors of extreme eigenvalues of matrices is that they provide low-rank approximations of a matrix. Recall

that if A is a symmetric matrix with eigenvalues $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$ and a corresponding orthonormal basis of column eigenvectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, then

$$A = \sum_i \alpha_i \boldsymbol{v}_i \boldsymbol{v}_i^T.$$

We can measure how well a matrix B approximates a matrix A by either the operator norm $\|A - B\|$ or the Frobenius norm $\|A - B\|_F$, where we recall

$$\|M\| \stackrel{\text{def}}{=} \max_{\boldsymbol{x}} \frac{\|M\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \quad \text{and} \quad \|M\|_F \stackrel{\text{def}}{=} \sqrt{\sum_{i,j} M(i,j)^2}.$$

Using the Courant-Fischer Theorem, one can prove that for every k, the best approximation of A by a rank-k matrix is given by summing the terms $\alpha_i \boldsymbol{v}_i \boldsymbol{v}_i^T$ over the k values of i for which $|\alpha_i|$ is largest. This holds regardless of whether we measure the quality of approximation in the operator or Frobenius norm.

When the difference between A and its best rank-k approximation is small, it explains why the eigenvectors of the largest k eigenvalues of A should provide a lot of information about A. However, one must be careful when applying this intuition as the analogous eigenvectors of the Laplacian correspond to is smallest eigenvalues. Perhpas the best way to explain the utility of these small eigenvectors is to observe that they provide the best low-rank approximation of the pseudoinverse of the Laplacian.

16.6 Elementary Facts

We list some elementary facts about the extreme eigenvalues of the Laplacian and adjacency matrices. We recommend deriving proofs yourself, or consulting the suggested references.

1. The all-1s vector is always an eigenvector of L_G of eigenvalue 0.

2. The largest eigenvalue of the adjacency matrix is at least the average degree of a vertex of G and at most the maximum degree of a vertex of G (see [9] or [10, Section 3.2]).

3. If G is connected, then $\alpha_1 > \alpha_2$ and the eigenvector of α_1 may be taken to be positive (this follows from the Perron-Frobenius theory; see [11]).

4. The all-1s vector is an eigenvector of A_G with eigenvalue α_1 if and only if G is an α_1-regular graph.

5. The multiplicity of 0 as an eigenvalue of L_G is equal to the number of connected components of L_G.

6. The largest eigenvalue of L_G is at most twice the maximum degree of a vertex in G.

7. $\alpha_n = -\alpha_1$ if and only if G is bipartite (see [12], or [10, Theorem 3.4]).

16.7 Spectral Graph Drawing

We can now explain the motivation behind Hall's spectral graph drawing technique [4]. Hall first considered the problem of assigning a real number $\boldsymbol{x}(a)$ to each vertex a so that $(\boldsymbol{x}(a) - \boldsymbol{x}(b))^2$ is small for most edges (a, b). This led him to consider the problem of minimizing (16.2). So as to avoid the degenerate solutions in which every vertex is mapped to zero, or any other value, he introduces the restriction that $\boldsymbol{x}$ be orthogonal to $\mathbf{b}1$. As the utility of the embedding does not really depend upon its scale, he suggested the normalization $\|\boldsymbol{x}\| = 1$. By the Courant-Fischer Theorem, the solution to the resulting optimization problem is precisely an eigenvector of the second-smallest eigenvalue of the Laplacian.

But, what if we want to assign the vertices to points in $\mathbb{R}^2$? The natural minimization problem,

$$\min_{\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^V} \sum_{(a,b) \in E} \|(\boldsymbol{x}(a), \boldsymbol{y}(a)) - (\boldsymbol{x}(b), \boldsymbol{y}(b))\|^2$$

such that

$$\sum_a (\boldsymbol{x}(a), \boldsymbol{y}(a)) = (0, 0)$$

typically results in the degenerate solution $\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{v}_2$. To ensure that the two coordinates are different, Hall introduced the restriction that $\boldsymbol{x}$ be orthogonal to $\boldsymbol{y}$. One can use the Courant-Fischer Theorem to show that the optimal solution is then given by setting $\boldsymbol{x} = \boldsymbol{v}_2$ and $\boldsymbol{y} = \boldsymbol{v}_3$, or by taking a rotation of this solution.

Hall observes that this embedding seems to cluster vertices that are close in the graph, and separate vertices that are far in the graph. For more sophisticated approaches to drawing graphs, we refer the reader to Chapter 15.

16.8 Algebraic Connectivity and Graph Partitioning

Many useful ideas in spectral graph theory have arisen from efforts to find quantitative analogs of qualitative statements. For example, it is easy to show

that $\lambda_2 > 0$ if and only if G is connected. This led Fiedler [13] to label λ_2 the *algebraic connectivity* of a graph, and to prove in various ways that better connected graphs have higher values of λ_2. This also led Fiedler to consider dividing the nodes of a graph into two pieces by choosing a real number t, and partitioning the nodes depending on whether or not $\boldsymbol{v}_2(a) \geq t$. For $t = 0$, this corresponds to selecting all vertices in the right-half of the spectral embedding of the graph.

```
S = find(v(:,2) >= 0);
plot(v(S,2),v(S,1),'o')
```
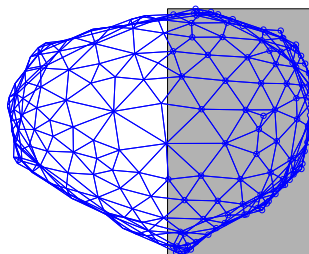
Fiedler proved [14] that for all $t \leq 0$, the set of nodes a for which $\boldsymbol{v}_2(a) \geq t$ forms a connected component. This type of "nodal domain theorem" was extended by van der Holst [15] to the set of a such that $\boldsymbol{v}(a) > 0$, when $\boldsymbol{v}$ is an eigenvector of λ_2 of minimal support.

The use of graph eigenvectors to partition graphs was also pioneered by Donath and Hoffman [16, 17] and Barnes [18]. It was popularized by experimental studies showing that it could give very good results [19, 20, 21, 22].

In many applications, one wants to partition the nodes of a graph into a few pieces of roughly equal size without removing too many edges (see Chapters 10 and 13). For simplicity, consider the problem of dividing the vertices of a graph into two pieces. In this case, we need merely identify one piece $S \subset V$. We then define $\partial(S)$ to be the set of edges with exactly one endpoint in S. We will also refer to S as a *cut*, as it implicitly divides the vertices into S and $V - S$, cutting all edges in $\partial(S)$. A tradeoff between the number of edges cut and the balance of the partition is obtained by dividing the first by a measure of the second, resulting in quantities called *cut ratio*, *sparsity*, *isoperimetric number*, and *conductance*, although these terms are sometimes used interchangeably. Wei and Cheng [23] suggested measuring the *ratio* of a cut, which they defined to be

$$R(S) \stackrel{\text{def}}{=} \frac{|\partial(S)|}{|S|\,|V - S|}.$$

Hagen and Kahng [24] observe that this quantity is always at least λ_2/n, and that $\boldsymbol{v}_2$ can be described as a relaxation of the characteristic vector[2] of the set S that minimizes $R(S)$.

Let $\boldsymbol{\chi}_S$ be the characteristic vector of a set S. For an unweighted graph G

---

[2]Here, we define the characteristic vector of a set to be the vector that is one at vertices inside the set and zero elsewhere.

we have

$$\boldsymbol{\chi}_S^T L_G \boldsymbol{\chi}_S = |\partial(S)|\,,$$

and

$$\sum_{a<b}(\boldsymbol{\chi}_S(a) - \boldsymbol{\chi}_S(b))^2 = |S|\,|V - S|\,.$$

So,

$$R(S) = \frac{\boldsymbol{\chi}_S^T L_G \boldsymbol{\chi}_S}{\sum_{a<b}(\boldsymbol{\chi}_S(a) - \boldsymbol{\chi}_S(b))^2}.$$

On the other hand, Fiedler [14] proved that

$$\lambda_2 = n \min_{\boldsymbol{x} \neq \mathbf{b}0} \frac{\boldsymbol{x}^T L_G \boldsymbol{x}}{\sum_{a<b}(\boldsymbol{x}(a) - \boldsymbol{x}(b))^2}.$$

If we impose the restriction that $\boldsymbol{x}$ be a zero-one valued vector and then minimize this last expression, we obtain the characteristic vector of the set of minimum ratio. As we have imposed a constraint on the vector $\boldsymbol{x}$, the minimum ratio obtained must be larger than λ_2. Hagen and Kahng make this observation, and suggest using $\boldsymbol{v}_2$ to try to find a set of low ratio by choosing some value t, and setting $S = \{a : \boldsymbol{v}(a) \geq t\}$.

One may actually prove that the set obtained in this fashion does not have ratio too much worse than the minimum. Statements of this form follow from discrete versions of Cheeger's inequality [25]. The cleanest version relates to the the *conductance* of a set S

$$\phi(S) \stackrel{\text{def}}{=} \frac{w(\partial(S))}{\min(\boldsymbol{d}(S), \boldsymbol{d}(V - S))},$$

where $\boldsymbol{d}(S)$ denotes the sum of the degrees of the vertices in S and $w(\partial(S))$ denotes the sum of the weights of the edges in $\partial(S)$. The conductance of the graph G is defined by

$$\phi_G = \min_{\emptyset \subset S \subset V} \phi(S).$$

By a similar relaxation argument, one can show

$$2\phi_G \geq \nu_2.$$

Sinclair and Jerrum's discrete version of Cheeger's inequality [26] says that

$$\nu_2 \leq \phi_G^2/2.$$

Moreover, their proof reveals that if $\boldsymbol{v}_2$ is an eigenvector of ν_2, then there exists a t so that

$$\phi\left(\left\{a : \boldsymbol{d}^{-1/2}(a)\boldsymbol{v}_2(a) \geq t\right\}\right) \leq \sqrt{2\nu_2}.$$

Other discretizations of Cheeger's inequality were proved around the same

time by a number of researchers. See [27, 28, 29, 30, 31]. We remark that Lawler and Sokal define conductance by

$$\frac{w(\partial(S))}{\boldsymbol{d}(S)\boldsymbol{d}(V-S)},$$

which is proportional to the *normalized cut* measure

$$\frac{w(\partial(S))}{\boldsymbol{d}(S)} + \frac{w(\partial(V-S))}{\boldsymbol{d}(V-S)}$$

popularized by Shi and Malik [21]. The advantage of this later formulation is that it has an obvious generalization to partitions into more than two pieces.

In general, the eigenvalues and entries of eigenvectors of Laplacian matrices will not be rational numbers; so, it is unreasonable to hope to compute them exactly. Mihail [32] proves that an approximation of the second-smallest eigenvector suffices. While her argument was stated for regular graphs, one can apply it to irregular, weighted graphs to show that for every vector $\boldsymbol{x}$ orthogonal to $\boldsymbol{d}^{1/2}$ there exists a t so that

$$\phi\left(\left\{a : \boldsymbol{d}^{-1/2}(a)\boldsymbol{x}(a) \geq t\right\}\right) \leq \sqrt{2\frac{\boldsymbol{x}^T N_G \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}}.$$

While spectral partitioning heuristics are easy to implement, they are neither the most effective in practice or in theory. Theoretically better algorithms have been obtained by linear programming [33] and by semi-definite programming [34]. Fast variants of these algorithms may be found in [35, 36, 37, 38, 39]. More practical algorithms are discussed in Chapters 10 and 13.

16.8.1　Convergence of Random Walks

If G is a connected, undirected graph, then the largest eigenvalue of W_G, ω_1, has multiplicity 1, equals 1, and has eigenvector $\boldsymbol{d}$. We may convert this eigenvector into a probability distribution $\boldsymbol{\pi}$ by setting

$$\boldsymbol{\pi} = \frac{\boldsymbol{d}}{\sum_a \boldsymbol{d}(a)}.$$

If $\omega_n \neq -1$, then the distribution of every random walk eventually converges to $\boldsymbol{\pi}$. The rate of this convergence is governed by how close $\max(\omega_2, -\omega_n)$ is to ω_1. For example, let $\boldsymbol{p}_t$ denote the distribution after t steps of a random walk that starts at vertex a. Then for every vertex b,

$$|\boldsymbol{p}_t(b) - \boldsymbol{\pi}(b)| \leq \sqrt{\frac{\boldsymbol{d}(b)}{\boldsymbol{d}(a)}} \left(1 - \max(\omega_2, -\omega_n)\right)^t.$$

One intuition behind Cheeger's inequality is that sets of small conductance are precisely the obstacles to the convergence of random walks.

For more information about random walks on graphs, we recommend the survey of Lovàsz [40] and the book by Doyle and Snell [2].

16.8.2 Expander Graphs

Some of the most fascinating graphs are those on which random walks mix quickly and which have high conductance. These are called *expander graphs*, and may be defined as the d-regular graphs for which all non-zero Laplacian eigenvalues are bounded away from zero. In the better expander graphs, all the Laplacian eigenvalues are close to d. One typically considers infinite families of such graphs in which d and a lower bound on the distance of the non-zero eigenvalues from d remain constant. These are counter-examples to many naive conjectures about graphs, and should be kept in mind whenever one is thinking about graphs. They have many amazing properties, and have been used throughout Theoretical Computer Science. In addition to playing a prominent role in countless theorems, they are used in the design of pseudo-random generators [41, 42, 43], error-correcting codes [44, 45, 46, 47, 48], fault-tolerant circuits [49] and routing networks [50].

The reason such graphs are called *expanders* is that all small sets of vertices in these graphs have unusually large numbers of neighbors. That is, their neighborhoods expand. For $S \subset V$, let $N(S)$ denote the set of vertices that are neighbors of vertices in S. Tanner [51] provides a lower bound on the size of $N(S)$ in bipartite graphs. In general graphs, it becomes the following.

Theorem 4 *Let $G = (V, E)$ be a d-regular graph on n vertices and set*

$$\epsilon = \max\left(1 - \frac{\lambda_2}{d}, \frac{\lambda_n}{d} - 1\right)$$

Then, for all $S \subseteq V$,

$$|N(S)| \geq \frac{|S|}{\epsilon^2(1 - \alpha) + \alpha},$$

where $|S| = \alpha n$.

The term ϵ is small when all of the eigenvalues are close to d. Note that when α is much less than ϵ^2, the term on the right is approximately $|S|/\epsilon^2$, which can be much larger than $|S|$.

An example of the pseudo-random properties of expander graphs is the "Expander Mixing Lemma". To understand it, consider choosing two subsets of the vertices, S and T of sizes αn and βn, at random. Let $\vec{E}(S, T)$ denote the set of ordered pairs (a, b) with $a \in S$, $b \in T$ and $(a, b) \in E$. The expected size of $\vec{E}(S, T)$ is $\alpha \beta d n$. This theorem tells us that for *every* pair of large sets S and T, the number of such pairs is approximately this quantity. Alternatively, one may view an expander as an approximation of the complete graph. The fraction of edges in the complete graph going from S to T is $\alpha \beta$. The following theorem says that the same is approximately true for all sufficiently large sets S and T.

Theorem 5 (Expander Mixing Lemma) *Let $G = (V, E)$ be a d-regular*

graph and set

$$\epsilon = \max\left(1 - \frac{\lambda_2}{d}, \frac{\lambda_n}{d} - 1\right)$$

Then, for every $S \subseteq V$ and $T \subseteq V$,

$$\left|\left|\vec{E}(S,T)\right| - \alpha\beta dn\right| \leq \epsilon dn\sqrt{(\alpha - \alpha^2)(\beta - \beta^2)},$$

where $|S| = \alpha n$ and $|T| = \beta n$.

This bound is a slight extension by Beigel, Margulis and Spielman [52] of a bound originally proved by Alon and Chung [53]. Observe that when α and β are greater than ϵ, the term on the right is less than $\alpha\beta dn$. Theorem 4 may be derived from Theorem 5.

We refer readers who would like to learn more about expander graphs to the survey of Hoory, Linial and Wigderson [54].

16.8.3   Ramanujan Graphs

Given the importance of λ_2, we should know how close it can be to d. Nilli [55] shows that it cannot be much closer than $2\sqrt{d-1}$.

Theorem 6 *Let G be an unweighted d-regular graph containing two edges (u_0, u_1) and (v_0, v_1) whose vertices are at distance at least $2k + 2$ from each other. Then*

$$\lambda_2 \leq d - 2\sqrt{d-1} + \frac{2\sqrt{d-1} - 1}{k+1}.$$

Amazingly, Margulis [56] and Lubotzky, Phillips and Sarnak [57] have constructed infinite families of d-regular graphs, called Ramanujan graphs, for which $\lambda_2 \geq d - 2\sqrt{d-1}$.

However, this is not the end of the story. Kahale [58] proves that vertex expansion by a factor greater than $d/2$ cannot be derived from bounds on λ_2. Expander graphs that have expansion greater than $d/2$ on small sets of vertices have been derived by Capalbo *et. al.* [59] through non-spectral arguments.

16.8.4   Bounding λ_2

I consider λ_2 to be the most interesting parameter of a connected graph. If it is large, the graph is an expander. If it is small, then the graph can be cut into two pieces without removing too many edges. Either way, we learn something about the graph. Thus, it is very interesting to find ways of estimating the value of λ_2 for families of graphs.

One way to explain the success of spectral partitioning heuristics is to prove that the graphs to which they are applied have small values of λ_2 or ν_2. A line of work in this direction was started by Spielman and Teng [60], who proved upper bounds on λ_2 for planar graphs and well-shaped finite element meshes.

Theorem 7 ([60]) *Let G be a planar graph with n vertices of maximum degree d, and let λ_2 be the second-smallest eigenvalue of its Laplacian. Then,*

$$\lambda_2 \leq \frac{8d}{n}.$$

This theorem has been extended to graphs of bounded genus by Kelner [61]. Entirely new techniques were developed by Biswal, Lee and Rao [62] to extend this bound to graphs excluding bounded minors. Bounds on higher Laplacian eigenvalues have been obtained by Kelner, Lee, Price and Teng [63].

Theorem 8 ([63]) *Let G be a graph with n vertices and constant maximum degree. If G is planar, has constant genus, or has a constant-sized forbidden minor, then*

$$\lambda_k \leq O(k/n).$$

Proving lower bounds on λ_2 is a more difficult problem. The dominant approach is to relate the graph under consideration to a graph with known eigenvalues, such as the complete graph. Write

$$L_G \succcurlyeq c L_H$$

if $L_G - c L_H \succcurlyeq 0$. In this case, we know that

$$\lambda_i(G) \geq c\lambda_i(H),$$

for all i. Inequalities of this form may be proved by identifying each edge of the graph H with a path in G. The resulting bounds are called Poincaré inequalities, and are closely related to the bounds used in the analysis of preconditioners in Chapter 12 and in related works [64, 65, 66, 67]. For examples of such arguments, we refer the reader to one of [68, 69, 70].

16.9   Coloring and Independent Sets

In the graph coloring problem one is asked to assign a color to every vertex of a graph so that every edge connects vertices of different colors, while using as few colors as possible. Replacing colors with numbers, we define a k-coloring of a graph $G = (V, E)$ to be a function $c : V \to \{1, \ldots, k\}$ such that

$$c(i) \neq c(j), \text{ for all } (i, j) \in E.$$

The *chromatic number* of a graph G, written $\chi(G)$, is the least k for which G has a k-coloring. Wilf [71] proved that the chromatic number of a graph may be bounded above by its largest adjacency eigenvalue.

Theorem 9 ([71])

$$\chi(G) \leq \alpha_1 + 1.$$

On the other hand, Hoffman [72] proved a lower bound on the chromatic number in terms of the adjacency matrix eigenvalues. When reading this theorem, recall that α_n is negative.

Theorem 10 *If G is a graph with at least one edge, then*

$$\chi(G) \geq \frac{\alpha_1 - \alpha_n}{-\alpha_n} = 1 + \frac{\alpha_1}{-\alpha_n}.$$

In fact, this theorem holds for arbitrary weighted graphs. Thus, one may prove lower bounds on the chromatic number of a graph by assigning a weight to every edge, and then computing the resulting ratio.

It follows from Theorem 10 that G is not bipartite if $|\alpha_n| < \alpha_1$. Moreover, as $|\alpha_n|$ becomes closer to 0, more colors are needed to properly color the graph. Another way to argue that graphs with small $|\alpha_n|$ are far from being bipartite was found by Trevisan [73]. To be precise, Trevisan proves a bound, analogous to Cheeger's inequality, relating $|E| - \max_{S \subset V} |\partial(S)|$ to the smallest eigenvalue of the *signless Laplacian matrix, $D_G + A_G$.*

An *independent set* of vertices in a graph G is a set $S \subseteq V$ such that no edge connects two vertices of S. The size of the largest independent set in a graph is called its *independence number*, and is denoted $\alpha(G)$. As all the nodes of one color in a coloring of G are independent, we know

$$\alpha(G) \geq n/\chi(G).$$

For regular graphs, Hoffman derived the following upper bound on the size of an independent set.

Theorem 11 *Let $G = (V, E)$ be a d-regular graph. Then*

$$\alpha(G) \leq n\frac{-\alpha_n}{d - \alpha_n}.$$

This implies Theorem 10 for regular graphs.

16.10 Perturbation Theory and Random Graphs

McSherry [74] observes that the spectral partitioning heuristics and the related spectral heuristics for graph coloring can be understood through matrix perturbation theory. For example, let G be a graph and let S be a subset

of the vertices of G. Without loss of generality, assume that S is the set of the first $|S|$ vertices of G. Then, we can write the adjacency matrix of G as

$$\begin{bmatrix} A(S) & 0 \\ 0 & A(V-S) \end{bmatrix} + \begin{bmatrix} 0 & A(S, V-S) \\ A(V-S, S) & 0 \end{bmatrix},$$

where we write $A(S)$ to denote the restriction of the adjacency matrix to the vertices in S, and $A(S, V-S)$ to capture the entries in rows indexed by S and columns indexed by $V-S$. The set S can be discovered from an examination of the eigenvectors of the left-hand matrix: it has one eigenvector that is positive on S and zero elsewhere, and another that is positive on $V-S$ and zero elsewhere. If the right-hand matrix is a "small" perturbation of the left-hand matrix, then we expect similar eigenvectors to exist in A. It seems reasonable that the right-hand matrix should be small if it contains few edges. Whether or not this may be made rigorous depends on the locations of the edges. We will explain McSherry's analysis, which makes this rigorous in certain random models.

We first recall the basics perturbation theory for matrices. Let A and B be symmetric matrices with eigenvalues $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$ and $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_n$, respectively. Let $M = A - B$. Weyl's Theorem, which follows from the Courant-Fischer Theorem, tells us that

$$|\alpha_i - \beta_i| \leq \|M\|$$

for all i. As M is symmetric, $\|M\|$ is merely the largest absolute value of an eigenvalue of M.

When some eigenvalue α_i is well-separated from the others, one can show that a small perturbation does not change the corresponding eigenvector too much. Demmel [75, Theorem 5.2] proves the following bound.

Theorem 12 *Let $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ be an orthonormal basis of eigenvectors of A corresponding to $\alpha_1, \ldots, \alpha_n$ and let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ be an orthonormal basis of eigenvectors of B corresponding to $\beta_1, \ldots, \beta_n$. Let θ_i be the angle between $\boldsymbol{v}_i$ and $\boldsymbol{w}_i$. Then,*

$$\frac{1}{2} \sin 2\theta_i \leq \frac{\|M\|}{\min_{j \neq i} |\alpha_i - \alpha_j|}.$$

McSherry applies these ideas from perturbation theory to analyze the behavior of spectral partitioning heuristics on random graphs that are generated to have good partitions. For example, he considered the planted partition model of Boppana [76]. This is defined by a weighted complete graph H determined by a $S \subset V$ in which

$$w(a, b) = \begin{cases} p & \text{if both or neither of } a \text{ and } b \text{ are in } S, \text{ and} \\ q & \text{if exactly one of } a \text{ and } b \text{ are in } S, \end{cases}$$

for $q < p$. A random unweighted graph G is then constructed by including

edge (a, b) in G with probability $w(a, b)$. For appropriate values of q and p, the cut determined by S is very likely to be the sparsest. If q is not too close to p, then the largest two eigenvalues of H are far from the rest, and correspond to the all-1s vector and a vector that is uniform and positive on S and uniform and negative on $V - S$. Using results from random matrix theory of Füredi and Komlós [77], Vu [78], and Alon, Krievlevich and Vu [79], McSherry proves that G is a slight perturbation of H, and that the eigenvectors of G can be used to recover the set S, with high probability.

Both McSherry [74] and Alon and Kahale [80] have shown that the eigenvectors of the smallest adjacency matrix eigenvalues may be used to k-color randomly generated k-colorable graphs. These graphs are generated by first partitioning the vertices into k sets, $S_1, \ldots, S_k$, and then adding edges between vertices in different sets with probability p, for some small p.

For more information on these and related results, we suggest the book by Kannan and Vempala [81].

16.11 Relative Spectral Graph Theory

Preconditioning (see Chapter 12) has inspired the study of the *relative eigenvalues of graphs*. These are the eigenvalues of $L_G L_H^+$, where L_G is the Laplacian of a graph G and L_H^+ is the pseudo-inverse of the Laplacian of a graph H. We recall that the pseudo-inverse of a symmetric matrix L is given by

$$\sum_{i : \lambda_i \neq 0} \frac{1}{\lambda_i} \boldsymbol{v}_i \boldsymbol{v}_i^T,$$

where the λ_i and $\boldsymbol{v}_i$ are the eigenvalues and eigenvectors of the matrix L. The eigenvalues of $L_G L_H^+$ reveal how well H approximates G.

Let K_n denote the complete graph on n vertices. All of the non-trivial eigenvalues of the Laplacian of K_n equal n. So, L_{K_n} acts as n times the identity on the space orthogonal to $\mathbf{b}1$. Thus, for every G the eigenvalues of $L_G L_{K_n}^+$ are just the eigenvalues of L_G divided by n, and the eigenvectors are the same. Many results on expander graphs, including those in Section 16.8.2, can be derived by using this perspective to treat an expander as an approximation of the complete graph (see [82]).

Recall that when L_G and L_H have the same range, $\kappa_f(L_G, L_H)$ is defined to be the largest non-zero eigenvalue of $L_G L_H^+$ divided by the smallest. The Ramanujan graphs are d-regular graphs G for which

$$\kappa_f(L_G, L_{K_n}) \leq \frac{d + 2\sqrt{d-1}}{d - 2\sqrt{d-1}}.$$

Batson, Spielman and Srivastava [82] prove that every graph H can be approximated by a sparse graph almost as well as this.

Theorem 13 *For every weighted graph G on n vertices and every $d > 1$, there exists a weighted graph H with at most $\lceil d(n-1) \rceil$ edges such that*

$$\kappa_f(L_G, L_H) \leq \frac{d + 1 + 2\sqrt{d}}{d + 1 - 2\sqrt{d}}.$$

Spielman and Srivastava [83] show that if one forms a graph H by sampling $O(n \log n/\epsilon^2)$ edges of G with probability proportional to their effective resistance and rescaling their weights, then with high probability $\kappa_f(L_G, L_H) \leq 1 + \epsilon$.

Spielman and Woo [84] have found a characterization of the well-studied *stretch* of a spanning tree with respect to a graph in terms of relative graph spectra. For simplicity, we just define it for unweighted graphs. If T is a spanning tree of a graph $G = (V, E)$, then for every $(a, b) \in E$ there is a unique path in T connecting a to b. The stretch of (a, b) with respect to T, written $\mathsf{st}_T(a, b)$, is the number of edges in that path in T. The stretch of G with respect to T is then defined to be

$$\mathsf{st}_T(G) \stackrel{\mathrm{def}}{=} \sum_{(a,b) \in E} \mathsf{st}_T(a, b).$$

Theorem 14 ([84])

$$\mathsf{st}_T(G) = \mathrm{trace}\left(L_G L_T^+\right).$$

See Chapter 12 for a proof.

16.12   Directed Graphs

There has been much less success in the study of the spectra of directed graphs, perhaps because the nonsymmetric matrices naturally associated with directed graphs are not necessarily diagonalizable. One naturally defines the adjacency matrix of a directed graph G by

$$A_G(a, b) = \begin{cases} 1 & \text{if } G \text{ has a directed edge from } b \text{ to } a \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, if we let $\boldsymbol{d}(a)$ denote the number of edges leaving vertex a and define D as before, then the matrix realizing the random walk on G is

$$W_G = A_G D_G^{-1}.$$

The Perron-Frobenius Theorem (see [11, 8]) tells us that if G is strongly connected, then A_G has a unique positive eigenvector $\boldsymbol{v}$ with a positive eigenvalue λ such that every other eigenvalue μ of A satisfies $|\mu| \leq \lambda$. The same holds for W_G. When $|\mu| < \lambda$ for all other eigenvalues μ, this vector is proportional to the unique limiting distribution of the random walk on G.

These Perron-Frobenius eigenvectors have proved incredibly useful in a number of situations. For instance, athey are at the heart of Google's PageRank algorithm for answering web search queries (see [85, 86]). This algorithm constructs a directed graph by associating vertices with web pages, and creating a directed edge for each link. It also adds a large number of low-weight edges by allowing the random walk to move to a random vertex with some small probability at each step. The PageRank score of a web page is then precisely the value of the Perron-Frobenius vector at the associated vertex. Interestingly, this idea was actually proposed by Bonacich [87, 88, 89] in the 1970's as a way of measuring the centrality of nodes in a social network. An analogous measure, using the adjacency matrix, was proposed by Berge [90, Chapter 4, Section 5] for ranking teams in sporting events. Palacios-Huerta and Volij [91] and Altman and Tennenholtz [92] have given abstract, axiomatic descriptions of the rankings produced by these vectors.

An related approach to obtaining rankings from directed graphs was proposed by Kleinberg [93]. He suggested using singular vectors of the directed adjacency matrix. Surprising, we are unaware of other combinatorially interesting uses of the singular values or vectors of matrices associated with directed graphs.

To avoid the complications of non-diagonalizable matrices, Chung [94] has defined a symmetric Laplacian matrix for directed graphs. Her definition is inspired by the observation that the degree matrix D used in the definition of the undirected Laplacian is the diagonal matrix of $\boldsymbol{d}$, which is proportional to the limiting distribution of a random walk on an undirected graph. Chung's Laplacian for directed graphs is constructed by replacing $\boldsymbol{d}$ by the Perron-Frobenius vector for the random walk on the graph. Using this Laplacian, she derives analogs of Cheeger's inequality, defining conductance by counting edges by the probability they appear in a random walk [95].

16.13   Concluding Remarks

Many fascinating and useful results in Spectral Graph Theory are omitted in this survey. For those who want to learn more, the following books and survey papers take an approach in the spirit of this Chapter: [96, 97, 98, 81, 3, 40]. I also recommend [10, 99, 6, 100, 101].

Anyone contemplating Spectral Graph Theory should be aware that there are graphs with very pathological spectra. Expanders could be considered ex-

amples. But, Strongly Regular Graphs (which only have 3 distinct eigenvalues) and Distance Regular Graphs should also be considered. Excellent treatments of these appear in some of the aforementioned works, and also in [6, 102].

Bibliography

[1] Bollobás, B., *Modern graph theory*, Springer-Verlag, New York, 1998.

[2] Doyle, P. G. and Snell, J. L., *Random Walks and Electric Networks*, Vol. 22 of *Carus Mathematical Monographs*, Mathematical Association of America, 1984.

[3] Chung, F. R. K., *Spectral Graph Theory*, American Mathematical Society, 1997.

[4] Hall, K. M., "An r-dimensional quadratic placement algorithm," *Management Science*, Vol. 17, 1970, pp. 219–229.

[5] Gould, S. H., *Variational Methods for Eigenvalue Problems*, Dover, 1995.

[6] Godsil, C., *Algebraic Combinatorics*, Chapman & Hall, 1993.

[7] van der Holst, H., Lovàsz, L., and Schrijver, A., "The Colin de Verdière Graph Parameter," *Bolyai Soc. Math. Stud.*, Vol. 7, 1999, pp. 29–85.

[8] Horn, R. A. and Johnson, C. R., *Matrix Analysis*, Cambridge University Press, 1985.

[9] Collatz, L. and Sinogowitz, U., "Spektren endlicher Grafen," *Abh. Math. Sem. Univ. Hamburg*, Vol. 21, 1957, pp. 63–77.

[10] Cvetković, D. M., Doob, M., and Sachs, H., *Spectra of Graphs*, Academic Press, 1978.

[11] Bapat, R. B. and Raghavan, T. E. S., *Nonnegative Matrices and Applications*, No. 64 in Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1997.

[12] Hoffman, A. J., "On the Polynomial of a Graph," *The American Mathematical Monthly*, Vol. 70, No. 1, 1963, pp. 30–36.

[13] Fiedler., M., "Algebraic connectivity of graphs," *Czechoslovak Mathematical Journal*, Vol. 23, No. 98, 1973, pp. 298–305.

[14] Fiedler, M., "A property of eigenvectors of nonnegative symmetric matrices and its applications to graph theory," *Czechoslovak Mathematical Journal*, Vol. 25, No. 100, 1975, pp. 618–633.

[15] van der Holst, H., "A Short Proof of the Planarity Characterization of Colin de Verdière," *Journal of Combinatorial Theory, Series B*, Vol. 65, No. 2, 1995, pp. 269 – 272.

[16] Donath, W. E. and Hoffman, A. J., "Algorithms for Partitioning Graphs and Computer Logic Based on Eigenvectors of Connection Matrices," *IBM Technical Disclosure Bulletin*, Vol. 15, No. 3, 1972, pp. 938–944.

[17] Donath, W. E. and Hoffman, A. J., "Lower Bounds for the Partitioning of Graphs," *IBM Journal of Research and Development*, Vol. 17, No. 5, Sept. 1973, pp. 420–425.

[18] Barnes, E. R., "An Algorithm for Partitioning the Nodes of a Graph," *SIAM Journal on Algebraic and Discrete Methods*, Vol. 3, No. 4, 1982, pp. 541–550.

[19] Simon, H. D., "Partitioning of Unstructured Problems for Parallel Processing," *Computing Systems in Engineering*, Vol. 2, 1991, pp. 135–148.

[20] Pothen, A., Simon, H. D., and Liou, K.-P., "Partitioning Sparse Matrices with Eigenvectors of Graphs," *SIAM Journal on Matrix Analysis and Applications*, Vol. 11, No. 3, 1990, pp. 430–452.

[21] Shi, J. B. and Malik, J., "Normalized Cuts and Image Segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 22, No. 8, Aug. 2000, pp. 888–905.

[22] Ng, A. Y., Jordan, M. I., and Weiss, Y., "On Spectral Clustering: Analysis and an algorithm," *Adv. in Neural Inf. Proc. Sys. 14*, 2001, pp. 849–856.

[23] Wei, Y.-C. and Cheng, C.-K., "Ratio cut partitioning for hierarchical designs," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 10, No. 7, Jul 1991, pp. 911–921.

[24] Hagen, L. and Kahng, A. B., "New Spectral Methods for Ratio Cut Partitioning and Clustering," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 11, 1992, pp. 1074–1085.

[25] Cheeger, J., "A lower bound for smallest eigenvalue of the Laplacian," *Problems in Analysis*, Princeton University Press, 1970, pp. 195–199.

[26] Sinclair, A. and Jerrum, M., "Approximate Counting, Uniform Generation and Rapidly Mixing Markov Chains," *Information and Computation*, Vol. 82, No. 1, July 1989, pp. 93–133.

[27] Lawler, G. F. and Sokal, A. D., "Bounds on the L^2 Spectrum for Markov Chains and Markov Processes: A Generalization of Cheeger's Inequality," *Transactions of the American Mathematical Society*, Vol. 309, No. 2, 1988, pp. 557–580.

[28] Alon, N. and Milman, V. D., "λ_1, Isoperimetric inequalities for graphs, and superconcentrators," *J. Comb. Theory, Ser. B*, Vol. 38, No. 1, 1985, pp. 73–88.

[29] Alon, N., "Eigenvalues and expanders," *Combinatorica*, Vol. 6, No. 2, 1986, pp. 83–96.

[30] Dodziuk, J., "Difference Equations, Isoperimetric Inequality and Transience of Certain Random Walks," *Transactions of the American Mathematical Society*, Vol. 284, No. 2, 1984, pp. 787–794.

[31] Varopoulos, N. T., "Isoperimetric inequalities and Markov chains," *Journal of Functional Analysis*, Vol. 63, No. 2, 1985, pp. 215 – 239.

[32] Mihail, M., "Conductance and Convergence of Markov Chains—A Combinatorial Treatment of Expanders," *30th Annual IEEE Symposium on Foundations of Computer Science*, 1989, pp. 526–531.

[33] Leighton, T. and Rao, S., "Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms," *Journal of the ACM*, Vol. 46, No. 6, Nov. 1999, pp. 787–832.

[34] Arora, S., Rao, S., and Vazirani, U., "Expander flows, geometric embeddings and graph partitioning," *J. ACM*, Vol. 56, No. 2, 2009, pp. 1–37.

[35] Khandekar, R., Rao, S., and Vazirani, U., "Graph partitioning using single commodity flows," *J. ACM*, Vol. 56, No. 4, 2009, pp. 1–15.

[36] Sherman, J., "Breaking the Multicommodity Flow Barrier for O(sqrt(log n))-approximations to Sparsest Cut," *Proceedings of the 50th IEEE Symposium on Foundations of Computer Science*, 2009, pp. 363 – 372.

[37] Arora, S., Hazan, E., and Kale, S., "O(sqrt (log n)) Approximation to Sparsest Cut in $\tilde{O}(n^2)$ Time," *45th IEEE Symposium on Foundations of Computer Science*, 2004, pp. 238–247.

[38] Arora, S. and Kale, S., "A combinatorial, primal-dual approach to semidefinite programs," *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, 2007, pp. 227–236.

[39] Orecchia, L., Schulman, L. J., Vazirani, U. V., and Vishnoi, N. K., "On partitioning graphs via single commodity flows," *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, 2008, pp. 461–470.

[40] Lovász, L., "Random walks on graphs: a survey," *Combinatorics, Paul Erdös is Eighty, Vol. 2*, edited by T. S. D. Miklos, V. T. Sos, Janos Bólyai Mathematical Society, Budapest, 1996, pp. 353–398.

[41] Impagliazzo, R. and Zuckerman, D., "How to recycle random bits," *30th annual IEEE Symposium on Foundations of Computer Science*, 1989, pp. 248–253.

[42] Karp, R. M., Pippenger, N., and Sipser, M., "A time randomness trade-off," *AMS Conf. on Probabilistic Computational Complexity*, Durham, New Hampshire, 1985.

[43] Ajtai, M., Komlós, J., and Szemerédi, E., "Deterministic Simulation in LOGSPACE," *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, 1987, pp. 132–140.

[44] Alon, N., Bruck, J., Naor, J., Naor, M., and Roth, R. M., "Construction of Asymptotically Good Low-Rate Error-Correcting Codes through Pseudo-Random Graphs," *IEEE Transactions on Information Theory*, Vol. 38, No. 2, March 1992, pp. 509–516.

[45] Sipser, M. and Spielman, D., "Expander codes," *IEEE Transactions on Information Theory*, Vol. 42, No. 6, Nov 1996, pp. 1710–1722.

[46] Zemor, G., "On expander codes," *IEEE Transactions on Information Theory*, Vol. 47, No. 2, Feb 2001, pp. 835–837.

[47] Barg, A. and Zemor, G., "Error exponents of expander codes," *IEEE Transactions on Information Theory*, Vol. 48, No. 6, Jun 2002, pp. 1725–1729.

[48] Guruswami, V., "Guest column: error-correcting codes and expander graphs," *SIGACT News*, Vol. 35, No. 3, 2004, pp. 25–41.

[49] Pippenger, N., "On networks of noisy gates," *Proceedings of the 26th Ann. IEEE Symposium on Foundations of Computer Science*, 1985, pp. 30–38.

[50] Pippenger, N., "Self-Routing Superconcentrators," *Journal of Computer and System Sciences*, Vol. 52, No. 1, Feb. 1996, pp. 53–60.

[51] Tanner, R. M., "Explicit construction of concentrators from generalized n-gons," *SIAM J. Algebraic Discrete Methods*, Vol. 5, 1984, pp. 287–293.

[52] Beigel, R., Margulis, G., and Spielman, D. A., "Fault diagnosis in a small constant number of parallel testing rounds," *SPAA '93: Proceedings of the fifth annual ACM symposium on Parallel algorithms and architectures*, ACM, New York, NY, USA, 1993, pp. 21–29.

[53] Alon, N. and Chung, F., "Explicit Construction of Linear Sized Tolerant Networks," *Discrete Mathematics*, Vol. 72, 1988, pp. 15–19.

[54] Hoory, S., Linial, N., and Wigderson, A., "Expander Graphs and Their Applications," *Bulletin of the American Mathematical Society*, Vol. 43, No. 4, 2006, pp. 439–561.

[55] Nilli, A., "On the second eigenvalue of a graph," *Discrete Math*, Vol. 91, 1991, pp. 207–210.

[56] Margulis, G. A., "Explicit group theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators," *Problems of Information Transmission*, Vol. 24, No. 1, July 1988, pp. 39–46.

[57] Lubotzky, A., Phillips, R., and Sarnak, P., "Ramanujan Graphs," *Combinatorica*, Vol. 8, No. 3, 1988, pp. 261–277.

[58] Kahale, N., "Eigenvalues and expansion of regular graphs," *J. ACM*, Vol. 42, No. 5, 1995, pp. 1091–1106.

[59] Capalbo, M., Reingold, O., Vadhan, S., and Wigderson, A., "Randomness conductors and constant-degree lossless expanders," *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, 2002, pp. 659–668.

[60] Spielman, D. A. and Teng, S.-H., "Spectral partitioning works: Planar graphs and finite element meshes," *Linear Algebra and its Applications*, Vol. 421, No. 2-3, 2007, pp. 284 – 305, Special Issue in honor of Miroslav Fiedler.

[61] Kelner, J. A., "Spectral partitioning, eigenvalue bounds, and circle packings for graphs of bounded genus," *SIAM J. Comput.*, Vol. 35, No. 4, 2006, pp. 882–902.

[62] Biswal, P., Lee, J., and Rao, S., "Eigenvalue bounds, spectral partitioning, and metrical deformations via flows," *Journal of the ACM*, 2010, to appear.

[63] Kelner, J. A., Lee, J., Price, G., and Teng, S.-H., "Higher Eigenvalues of Graphs," *Proceedings of the 50th IEEE Symposium on Foundations of Computer Science*, 2009, pp. 735–744.

[64] Vaidya, P. M., "Solving linear equations with symmetric diagonally dominant matrices by constructing good preconditioners." Unpublished manuscript UIUC 1990. A talk based on the manuscript was presented at the IMA Workshop on Graph Theory and Sparse Matrix Computation, October 1991, Minneapolis.

[65] Bern, M., Gilbert, J., Hendrickson, B., Nguyen, N., and Toledo, S., "Support-graph preconditioners," *SIAM Journal on Matrix Analysis and Applications*, Vol. 27, No. 4, 2006, pp. 930–951.

[66] Boman, E. G. and Hendrickson, B., "Support Theory for Preconditioning," *SIAM Journal on Matrix Analysis and Applications*, Vol. 25, No. 3, 2003, pp. 694–717.

[67] Spielman, D. A. and Teng, S.-H., "Nearly-Linear Time Algorithms for Preconditioning and Solving Symmetric, Diagonally Dominant

Linear Systems," *CoRR*, Vol. abs/cs/0607105, 2009, Available at
`http://www.arxiv.org/abs/cs.NA/0607105`.

[68] Diaconis, P. and Stroock, D., "Geometric Bounds for Eigenvalues of
Markov Chains," *The Annals of Applied Probability*, Vol. 1, No. 1, 1991,
pp. 36–61.

[69] Guattery, S., Leighton, T., and Miller, G. L., "The Path Resistance
Method for Bounding the Smallest Nontrivial Eigenvalue of a Lapla-
cian," *Combinatorics, Probability and Computing*, Vol. 8, 1999, pp. 441–
460.

[70] Guattery, S. and Miller, G. L., "Graph Embeddings and Laplacian
Eigenvalues," *SIAM Journal on Matrix Analysis and Applications*,
Vol. 21, No. 3, 2000, pp. 703–723.

[71] Wilf, H. S., "The Eigenvalues of a Graph and its Chromatic Number,"
J. London math. Soc., Vol. 42, 1967, pp. 330–332.

[72] Hoffman, A. J., "On eigenvalues and colorings of graphs," *Graph Theory
and its Applications*, Academic Press, New York, 1970, pp. 79–92.

[73] Trevisan, L., "Max cut and the smallest eigenvalue," *Proceedings of the
41st Annual ACM Symposium on Theory of Computing*, 2009, pp. 263–
272.

[74] McSherry, F., "Spectral Partitioning of Random Graphs," *Proceedings
of the 42nd IEEE Symposium on Foundations of Computer Science*,
2001, pp. 529–537.

[75] Demmel, J., *Applied Numerical Linear Algebra*, SIAM, 1997.

[76] Boppana, R. B., "Eigenvalues and graph bisection: an average-case anal-
ysis," *Proc. 28th IEEE Symposium on Foundations of Computer Sci-
ence*, 1987, pp. 280–285.

[77] Füredi, Z. and Komlós, J., "The eigenvalues of random symmetric ma-
trices," *Combinatorica*, Vol. 1, No. 3, 1981, pp. 233–241.

[78] Vu, V., "Spectral Norm of Random Matrices," *Combinatorica*, Vol. 27,
No. 6, 2007, pp. 721–736.

[79] Alon, N., Krivelevich, M., and Vu, V. H., "On the concentration of eigen-
values of random symmetric matrices," *Israel Journal of Mathematics*,
Vol. 131, No. 1, 2002, pp. 259–267.

[80] Alon, N. and Kahale, N., "A Spectral Technique for Coloring Ran-
dom 3-Colorable Graphs," *SIAM Journal on Computing*, Vol. 26, 1997,
pp. 1733–1748.

[81] Kannan, R. and Vempala, S., "Spectral Algorithms," *Foundations and Trends in Theoretical Computer Science*, Vol. 4, No. 3-4, 2008, pp. 132–288.

[82] Batson, J. D., Spielman, D. A., and Srivastava, N., "Twice-Ramanujan sparsifiers," *Proceedings of the 41st Annual ACM Symposium on Theory of computing*, 2009, pp. 255–262.

[83] Spielman, D. A. and Srivastava, N., "Graph sparsification by effective resistances," *SIAM Journal on Computing*, 2010, To appear. Preliminary version appeared in the Proceedings of the 40th Annual ACM Symposium on Theory of Computing.

[84] Spielman, D. A. and Woo, J., "A Note on Preconditioning by Low-Stretch Spanning Trees," *CoRR*, Vol. abs/0903.2816, 2009, Available at http://arxiv.org/abs/0903.2816.

[85] Page, L., Brin, S., Motwani, R., and Winograd, T., "The PageRank Citation Ranking: Bringing Order to the Web," Tech. rep., Computer Science Department, Stanford University, 1998.

[86] Langville, A. N. and Meyer, C. D., *Google's PageRank and Beyond: The Science of Search Engine Rankings*, Princeton University Press, Princeton, NJ, USA, 2006.

[87] Bonacich, P., "Technique for Analyzing Overlapping Memberships," *Sociological Methodology*, Vol. 4, 1972, pp. 176–185.

[88] Bonacich, P., "Factoring and Weighting Approaches to Status Scores and Clique Identification," *Journal of Mathematical Sociology*, Vol. 2, 1972, pp. 113–120.

[89] Bonacich, P., "Power and Centrality: A family of Measures," *American Journal of Sociology*, Vol. 92, 1987, pp. 1170–1182.

[90] Berge, C., *Graphs*, North-Holland, 1985.

[91] Palacios-Huerta, I. and Volij, O., "The Measurement of Intellectual Influence," *Econometrica*, Vol. 72, No. 3, 2004, pp. 963–977.

[92] Altman, A. and Tennenholtz, M., "Ranking systems: the PageRank axioms," *Proceedings of the 6th ACM Conference on Electronic Commerce*, 2005, pp. 1–8.

[93] Kleinberg, J., "Authoratitive sources in a hyperlinked environment," *Journal of the ACM*, Vol. 48, 1999, pp. 604–632.

[94] Chung, F., "The Diameter and Laplacian Eigenvalues of Directed Graphs," *The Electronic Journal of Combinatorics*, Vol. 13, No. 1, 2006.

[95] Chung, F., "Laplacians and the Cheeger Inequality for Directed Graphs," *Annals of Combinatorics*, Vol. 9, No. 1, 2005, pp. 1–19.

[96] Mohar, B. and Poljak, S., "Eigenvalues in combinatorial optimization," *Combinatorial and graph-theoretical problems in linear algebra*, IMA Volumes in Mathematics and Its Applications, Springer-Verlag, 1993, pp. 107–151.

[97] Mohar, B., "The Laplacian spectrum of graphs," *Graph Theory, Combinatorics, and Applications*, Wiley, 1991, pp. 871–898.

[98] Brouwer, A. E. and Haemers, W. H., "Spectra of Graphs," Electronic Book. Available at `http://www.win.tue.nl/~aeb/ipm.pdf`.

[99] Godsil, C. and Royle, G., *Algebraic Graph Theory*, Graduate Texts in Mathematics, Springer, 2001.

[100] Biggs, N. L., *Algebraic graph theory*, Cambridge Tracts in Math., Cambridge University Press, London, New York, 1974.

[101] Brualdi, R. A. and Ryser, H. J., *Combinatorial Matrix Theory*, Cambridge University Press, New York, 1991.

[102] Brouwer, A. E., Cohen, A. M., and Neumaier, A., *Distance-Regular Graphs.*, Springer Verlag, 1989.