Notes 09: Graph spectrum and Laplacian

We will consider undirected graph (without parallel edges or self-loops). We sometimes allow positive edge weights.

1. Graph spectrum

Consider an unweighted graph G on vertex set V of size n. its adjacency matrix A is a V-by-V matrix (i.e. rows are indexed by V, and so are the columns), whose V_{ij}-entry A_{ij} is 1 if (i,j) is an edge in G, and A_{ij} is 0 otherwise.

More generally, given a weighted graph with edge weights w_{ij} for edge (i,j), the ij-entry of the adjacency matrix A_{ij} is simply w_{ij}. Non-edges have weight 0 by convention.

Since we only consider undirected graph, its adjacency matrix A is symmetric. So A has n real eigenvalues and eigenvectors.

1.1. Complete graph.

For the complete graph, the adjacency matrix $A = J - I$, where J is the all-one matrix.

$J = \mathbb{1}\mathbb{1}^\top$, so it has an eigenvalue n (with eigenvector $\mathbb{1}$), and eigenvalue 0 with multiplicity $n-1$. Here $\mathbb{1}$ denotes the all-one vector (in $\mathbb{R}^V$).

The identity matrix I has eigenvalue 1 with multiplicity n.

Subtracting I from J decreases all eigenvalues by 1. So A has eigenvalue $n-1$ (with eigenvector 1) and eigenvalue -1 with multiplicity $n-1$.

2. Degree

The degree of a vertex of an unweighted graph is the number of edges incident to it. Let Δ denote the maximum degree over all vertices.

Claim 2.1. *Any eigenvalues of the adjacency matrix A is at most Δ in magnitude.*

Proof. Let v be an eigenvector of A with eigenvalue λ, so that $Av = \lambda v$. Assume i is the vertex that maximizes $|v_i|$. Then

$$|(\lambda v)_i| = |(Av)_i| = \left|\sum_{j\in V} A_{ij} v_j\right| \leqslant \sum_{j\sim i} |v_j| \leqslant \Delta |v_i|. \qquad \square$$

3. Graph Laplacian

Given an edge $e = (i,j)$ of weight w_e, its Laplacian is $L_e = w_e(\mathbb{1}_i - \mathbb{1}_j)(\mathbb{1}_1 - \mathbb{1}_j)^\top$. Here $\mathbb{1}_i \in \mathbb{R}^V$ denotes the vector that has entry 1 at vertex i and 0 elsewhere. This matrix has only four nonzero entries: w_e at (i,i) and (j,j), and $-w_e$ at (i,j) and (j,i).

Given a weighted graph G, its Laplacian is

$$L_G = \sum_{e\in G} L_e = \sum_{e\in G} w_e(\mathbb{1}_i - \mathbb{1}_j)(\mathbb{1}_i - \mathbb{1}_j)^\top.$$

Note that the Laplacian of any edge is positive semidefinite (using our assumption that edge weights are nonnegative). Therefore the Laplacian of any graph is also positive semidefinite.

If we sort all its eigenvalues $\lambda_1 \leqslant \ldots \leqslant \lambda_n$, we know $0 \leqslant \lambda_1$. In fact the Laplacian always has eigenvalue 0, with eigenvector $\mathbb{1}$. What can we say about the second smallest eigenvalue λ_2?

4. Connectedness

Claim 4.1. *A graph is connected if and only if* 0 *is an eigenvalue of L_G with multiplicity* 1. *(i.e. $\lambda_2 > 0$)*

Proof. If G is disconnected, then V can be partitioned into two vertex subsets U and $\overline{U}$ with no edge between them. Then $\mathbb{1}_U$ (the vector that is all-one on U and all-zero outside U) and $\mathbb{1}_{\overline{U}}$ are linearly independent eigenvectors of L_G, both corresponding to eigenvalue 0.

Suppose G is connected and x is an eigenvector of L_G of eigenvalue 0, then $Lx = 0$ and $x^\top Lx = 0$. Last equality means $\sum_{(i,j)\in G} w_{ij}(x_i - x_j)^2 = 0$.

Since all edge weights are positive and $(x_i - x_j)^2$ is nonnegative, the sum can be zero only when all terms $w_{ij}(x_i - x_j)^2$ are zero. Given a vertex i, all neighbors j of i must have $x_j = x_i$, or else the sum is strictly positive. The same argument shows all vertices j in the same connected component as i have $x_j = x_i$. Since G is connected, we have $x_j = x_i$ for all vertices j, so v must be of the form $\alpha\mathbb{1}$ for some real number α. This means the eigenvalue 0 of L_G has multiplicity 1. □

5. Rayleigh quotient

We will relate eigenvalues and eigenvectors to optimization using Rayleigh quotient.

Definition 5.1. Given a real symmetric matrix B and a non-zero vector x, its Rayleigh quotient is $\dfrac{x^\top Bx}{x^\top x} = \dfrac{\sum_{ij} B_{ij}x_i x_j}{\sum_i x_i^2}$.

Let $\mu_1 \geqslant \ldots \geqslant \mu_n$ be the eigenvalues of a real symmetric matrix B with orthonormal eigenvectors $v_1, \ldots, v_n$, so that $B = \sum_{1\leqslant i\leqslant n} \mu_i v_i v_i^\top$.

Claim 5.2. $\mu_1 = \max_{x\neq 0} \frac{x^\top Bx}{x^\top x}$.

Proof. Express $x = c_1 v_1 + \cdots + c_n v_n$ as a linear combination in the eigenbasis. Then

$$x^\top Bx = \left(\sum_i c_i v_i^\top\right)\left(\sum_i \mu_i v_i v_i^\top\right)\left(\sum_i c_i v_i\right) = \sum_i \mu_i c_i^2$$

because v_i are orthonormal. Similarly,

$$x^\top x = \left(\sum_i c_i v_i^\top\right)\left(\sum_i c_i v_i\right) = \sum_i c_i^2 \ .$$

We have $x^\top Bx = \sum_i \mu_i c_i^2 \leqslant \mu_1 \sum_i c_i^2 = \mu_1 x^\top x$.

Equality can be achieved when $c_1 = 1$ and $c_2 = \cdots = c_n = 0$, i.e. x is the top eigenvector v_1. □

6. Courant-Fischer

Let B be a real symmetric matrix as in the previous section, with eigenvalues $\mu_1 \geqslant \ldots \geqslant \mu_n$ and orthonormal eigenvectors $v_1, \ldots, v_n$.

The following theorem says that the k-th largest eigenvalue is the answer to the following optimization problem: Among all k-dimensional subspaces S of $\mathbb{R}^n$, if we find the minimum Rayleigh quotient within S, which subspace S has the largest minimum Rayleigh quotient?

Theorem 6.1 (Courant–Fischer)**.**

$$\mu_k = \max_{\substack{\text{subspace } S\subseteq\mathbb{R}^n\\ \dim S = k}} \min_{\substack{x\in S\\ x\neq 0}} \frac{x^\top Bx}{x^\top x} = \min_{\substack{\text{subspace } S\subseteq\mathbb{R}^n\\ \dim S = n-k+1}} \max_{\substack{x\in S\\ x\neq 0}} \frac{x^\top Bx}{x^\top x}$$

Proof. We only prove the max-min term, since the min-max is similar.

Let S_k be the k-dimensional subspace spanned by the first k eigenvectors $v_1, \ldots, v_k$. Any $x \in S_k$, when expressed in the eigenbasis $x = \sum_i c_i v_i$, has $c_{k+1} = \cdots = c_n = 0$.

So $x^\top Bx = \sum_i \mu_i c_i^2 = \sum_{i\leqslant k} \mu_i c_i^2 \geqslant \mu_k \sum_{i\leqslant k} c_i^2 = \mu_k x^\top x$.

That means the minimum Rayleigh quotient over S_k is at least μ_k. Hence μ_k is at most the max-min.

To show μ_k is at least the max-min, note that any k-dimensional subspace must intersect the $n-k+1$ dimensional subspace T_k spanned by the bottom eigenvectors $v_{k+1}, \ldots, v_n$.

For any $x \in T_k \setminus \{0\}$, its Rayleigh quotient is at most μ_k because

$$x^\top Bx = \sum_i \mu_i c_i^2 = \sum_{i>k} \mu_i c_i^2 \leqslant \mu_k \sum_{i>k} c_i^2 = \mu_k x^\top x \ .$$

So the minimum Rayleigh quotient over any k-dimensional subspace S must be at most μ_k. And this remains true after maximizing over all subspaces S. □

7. Eigenvalue interlacing

Let B be a real symmetric matrix with eigenvalues $\beta_1 \geqslant \ldots \geqslant \beta_n$ in non-increasing order. Let C be a principle submatrix of B, obtained by deleting the same row and column. Let $\gamma_1 \geqslant \ldots \geqslant \gamma_{n-1}$ be its eigenvalues.

Theorem 7.1. $\beta_1 \geqslant \gamma_1 \geqslant \beta_2 \geqslant \gamma_2 \geqslant \ldots \geqslant \beta_{n-1} \geqslant \gamma_{n-1} \geqslant \beta_n$.

Proof. $\beta_k \geqslant \gamma_k$: By Courant–Fischer, β_k is maximizing a quantity (minimum Rayleigh quotient) over all k-dimensional subspace S. When restricting to only those subspaces S orthogonal to $\mathbb{1}_i$ (where i is the common index of the removed row and column), the maximum can only decrease. And for any vector x in a subspace S orthogonal to $\mathbb{1}_i$, its Rayleigh quotient wrt to B is the same as that wrt to C. Now the maximum becomes γ_k, again by Courant–Fischer. In summary,

$$\beta_k = \max_{\substack{S\subseteq\mathbb{R}^n \\ \dim S=k}} \min_{\substack{x\in S\\ x\neq 0}} \frac{x^\top Bx}{x^\top x} \geqslant \max_{\substack{S\subseteq\mathbb{R}^n \\ \dim S=k \\ S\perp \mathbb{1}_i}} \min_{\substack{x\in S\\ x\neq 0}} \frac{x^\top Bx}{x^\top x} = \max_{\substack{S\subseteq\mathbb{R}^{n-1} \\ \dim S=k}} \min_{\substack{x\in S\\ x\neq 0}} \frac{x^\top Cx}{x^\top x} = \gamma_k \ .$$

$\gamma_k \geqslant \beta_{k+1}$: By Courant–Fischer, β_{k+1} is maximizing a quantity (minimum Rayleigh quotient) over all $k+1$ dimensional subspaces S. Let S_* be the maximizing subspace. Since S_* has dimension $k+1$, when intersecting with the orthogonal complement of $\mathbb{1}_i$ (where i is the common index of the removed row and column), the intersection S' must have dimension at least k. So

$$\beta_{k+1} = \min_{s\in S_*} \frac{x^\top Bx}{x^\top x} \leqslant \min_{x\in S'\setminus\{0\}} \frac{x^\top Bx}{x^\top x} = \min_{x\in S'\setminus\{0\}} \frac{x^\top Cx}{x^\top x} \leqslant \max_{\dim S=k} \min_{x\in S\setminus\{0\}} \frac{x^\top Cx}{x^\top x} = \gamma_k \ . \qquad \square$$

Corollary 7.2. *Let B be any real symmetric matrix and C be a principle submatrix of B, obtained by removing the same set of k rows and columns from B. Call $\beta_1 \geqslant \ldots \geqslant \beta_n$ eigenvalues of B and $\gamma_1 \geqslant \ldots \geqslant \gamma_{n-k}$ eigenvalues of C. Then for any $1 \leqslant i \leqslant n-k$,*

$$\beta_i \geqslant \gamma_i \geqslant \beta_{i+k}.$$

8. Induced subgraphs

Sensitivity Conjecture was a well-known conjecture in Theoretical Computer Science that was open for more than 30 years. It is known to follow from another conjecture about graphs that was recently proved, with a remarkably short proof.

The d-dimensional hypercube G is the graph on 2^d vertices, labelled with all binary strings of length d, and two vertices are adjacent if their corresponding strings differ in exactly one position.

Theorem 8.1 (Huang)**.** *Any induced subgraph of the d-dimensional hypercube G that contains strictly more than half of the vertices of G must have maximum degree at least $\sqrt{d}$.*

Proof. Inductively define matrices

$$A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad A_d = \begin{pmatrix} A_{d-1} & I \\ I & -A_{d-1} \end{pmatrix}.$$

It can be checked that A_d is an ± 1-signed adjacency matrix of the d-dimensional hypercube, in the sense that an edge correspond to ± 1 entry, and zero otherwise.

Lemma 8.2. *A_d is a 2^d-by-2^d matrix whose eigenvalues are $\sqrt{d}$ with multiplicity 2^{d-1} and $-\sqrt{d}$ with multiplicity 2^{d-1}.*

Proof. It can be proved by induction that $A_d^2 = dI$, because for $d = 1$, $A_1^2 = I$, and in general

$$A_d^2 = \begin{pmatrix} A_{d-1}^2 + I & 0 \\ 0 & A_{d-1}^2 + I \end{pmatrix} = dI,$$

where the last equality uses the induction hypothesis.

Eigenvalues of the square of a real symmetric matrix are precisely the squares of its eigenvalues. Therefore all the eigenvalues of A_d are either $\sqrt{d}$ or $-\sqrt{d}$. Since trace of A_d is 0, and trace is well known to be equal to the sum of eigenvalues, exactly half of the eigenvalues of A_d are $\sqrt{d}$ and the rest $-\sqrt{d}$. $\square$

Any induced subgraph H of G such that H contains $2^{d-1}+1$ vertices (i.e. strictly more than half) corresponds to a principle submatrix C of A_d. By interlacing theorem, the top eigenvalue of C is at least the 2^{d-1}-th largest eigenvalue of A_d, which is $\sqrt{d}$.

The main result follows by applying Claim 2.1 about top eigenvalue and maximum degree, and noting that the proof of the claim still holds for ± 1-signed adjacency matrix. $\square$