Spectral Graph Theory

Vignesh M Pai

Contents

1 Introduction

1.1 Eigenvalues and Optimization

Let M be a n dimensional symmetric matrix.

Definition 1.1.1: Rayleigh Quotient

The Rayleigh quotient for a matrix M is defined as

$$R(\boldsymbol{x}, M) := \frac{\boldsymbol{x}^T M \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}$$

the matrix is ommitted if obvious from context.

Theorem 1.1.1

Let

$$\boldsymbol{x} \in \arg \max_{\boldsymbol{x} \in \mathbb{R}^n - \{0\}} R(\boldsymbol{x})$$

Such an $\boldsymbol{x}$ exists and is an eigenvector of M with the maximum eigenvalue μ_1. We can write a similar statement for the minimum eigenvalue by minimizing R.

Theorem 1.1.2: Spectral Theorem for Symmetric Matrices

There exist numbers $\mu_1, ..., \mu_n$ and orthonormal vectors $\boldsymbol{\psi}_1, ..., \boldsymbol{\psi}_n$ such that $M\boldsymbol{\psi}_i = \mu_i \boldsymbol{\psi}_i$ iff for $1 \leq i \leq n$

$$\boldsymbol{\psi}_i \in \arg \max_{\substack{\|\boldsymbol{x}\|=1 \\ \boldsymbol{x}^T \boldsymbol{\psi}_j = 0,\ j<i}} R(\boldsymbol{x})$$

or equivalently

$$\boldsymbol{\psi}_i \in \arg \min_{\substack{\|\boldsymbol{x}\|=1 \\ \boldsymbol{x}^T \boldsymbol{\psi}_j = 0,\ j>i}} R(\boldsymbol{x})$$

Theorem 1.1.3: Courant-Fischer Theorem

Let M have eigenvalues $\mu_1 \geq \mu_2 \geq ... \geq \mu_n$, then

$$\mu_k = \max_{\substack{S \subset \mathbb{R}^n \\ \dim(S)=k}} \min_{\substack{\boldsymbol{x} \in S \\ \boldsymbol{x} \neq 0}} R(\boldsymbol{x}) = \min_{\substack{T \subset \mathbb{R}^n \\ \dim(T)=n-k+1}} \max_{\substack{\boldsymbol{x} \in T \\ \boldsymbol{x} \neq 0}} R(\boldsymbol{x})$$

where S, T are subspaces of $\mathbb{R}^n$.

Theorem 1.1.4: Cauchy's Interlacing Theorem

Let A be a symmetric real matrix of dimension n. Let B be obtained by deleting the same row and column of A (N is a principal submatrix of dimension $n-1$). Let $\alpha_1 \geq ... \geq \alpha_n$ be the eigenvalues of A and $\beta_1 \geq ... \geq \beta_{n-1}$ be the eigenvalues of B. Then for $1 \leq k \leq n-1$

$$\alpha_k \geq \beta_k \geq \alpha_{k+1}$$

1.2 The Laplacian and Graph Drawing

Vectors on graphs are functions $V \to \mathbb{R}$, the vector $\mathbf{1}$ denotes the function $\mathbf{1}(a) = 1$. The degree function $\boldsymbol{d}$ is also a vector.

Matrices on graphs are functions $V \times V \to \mathbb{R}$ or can be viewed as linear operators on the space of vectors on graphs. Let G be a graph, then we use M_G to denote the adjacency matrix, D_G to denote the diagonal matrix of vertex degrees, and $L_G = D_G - M_G$ to denote the Laplacian matrix. Observe that $M_G \mathbf{1} = \boldsymbol{d}$ and $L_G \mathbf{1} = \mathbf{0}$.

The Laplacian is also a natural quadratic form on a graph

$$\boldsymbol{x}^T L_G \boldsymbol{x} = \sum_{(a,b) \in E(G)} w_{a,b} (\boldsymbol{x}(a) - \boldsymbol{x}(b))^2$$

Let $\lambda_1 = 0 \leq \lambda_2 \leq ... \leq \lambda_n$ be the eigenvalues of L_G with eigenvectors $\boldsymbol{\psi}_1, ..., \boldsymbol{\psi}_n$.

Lemma 1.2.1

G is connected iff $\lambda_2 \neq 0$.

We draw a graph in k dimensions by using the eigenvectors corresponding to $\lambda_2, ..., \lambda_{k+1}$ as the coordinates of vertices. These coordinates minimize the expression (excluding coordinates that lead to trivial drawings):

$$\sum_{(a,b) \in G} w_{a,b} \|x(a) - x(b)\|^2 = \sum_{i=1}^{k} \boldsymbol{x_i}^T L_G \boldsymbol{x_i}$$

where $x : V \to \mathbb{R}^k, x = (\boldsymbol{x_1}, ..., \boldsymbol{x_k})$ is the coordinate function.

Theorem 1.2.2: Hall's Drawing Theorem

Let $\boldsymbol{x_1}, ..., \boldsymbol{x_k}$ be orthonormal vectors that are orthogonal to $\mathbf{1}$, then

$$\sum_{i=1}^{k} \boldsymbol{x_i}^T L_G \boldsymbol{x_i} \geq \sum_{i=2}^{k+1} \lambda_1$$

where equality holds for $\boldsymbol{\psi}_j^T \boldsymbol{x_i} = 0$ for all i and $j > k+1$.

1.3 Adjacency Matrices

Let the eigenvalues of M_G be $\mu_1 \geq ... \geq \mu_n$.

Lemma 1.3.1

Let d_{avg} and d_{max} be the average and maximum degrees respectively, then

$$d_{avg} \leq \mu_1 \leq d_{max}$$

further, if H be a subgraph of G, then

$$d_{avg}(H) \leq \mu_1$$

Lemma 1.3.2

If G is connected and $\mu_1 = d_{max}$ then G is d_{max}-regular.

Theorem 1.3.3: Wilf's Theorem

Let $\chi(G)$ be the chromatic number of G, then

$$\chi(G) \leq \lfloor \mu_1 \rfloor + 1$$

Lemma 1.3.4

Let G be a connected weighted graph and $\boldsymbol{\phi}$ be a non negative eigenvector of M_G, then $\boldsymbol{\phi}$ is strictly positive.

Theorem 1.3.5: Perron-Frobenius Theorem

Let G be connected, then

1. μ_1 has a strictly positive eigenvector
2. $\mu_1 \geq -\mu_n$
3. $\mu_1 > \mu_2$

Lemma 1.3.6

If G is bipartite, then the eigenvalues of M_G are symmetric about 0.

Theorem 1.3.7

Let G be connected, $\mu_1 = -\mu_n$ iff G is bipartite.

1.4 Comparing Graphs

We introduce the partial order $\succeq$ on matrices as

$$A \succeq B \iff \forall \boldsymbol{x}, \boldsymbol{x}^T A \boldsymbol{x} \geq \boldsymbol{x}^T B \boldsymbol{x}$$

In particular $A \succeq 0$ means A is positive semidefinite. For graphs G, H on the same set of vertices, we write $G \succeq H$ iff $L_G \succeq L_H$. If H is a subset of G, we have

$$G \succeq H$$

For a graph H, define $c \cdot H$ to be the graph H with each edge weight multiplied by c. Let $\lambda_k(H)$ denote the kth smallest eigenvalue of L_H.

Lemma 1.4.1

If G, H are graphs

$$G \succeq c \cdot H \implies \lambda_k(G) \geq c\lambda_k(H)$$

Let $G_{a,b}$ be the graph with only the edge (a, b). The proof of the following lemmas follow trivially from the Cauchy Schwarz inequality applied to the Laplacian.

Lemma 1.4.2

Let P_n be the path graph on n vertices between vertex 1 and n.

$$G_{1,n} \preceq (n-1)P_n$$

Lemma 1.4.3: Extension to Weighted Paths

Let $P_{n,w}$ be the weighted path graph on n vertices with w_i the weight on the edge $(i, i+1)$.

$$G_{1,n} \preceq \left(\sum_{i=1}^{n-1} \frac{1}{w_i}\right) P_{n,w}$$

We use these lemmas to prove bounds on the eigenvalues of graphs.

Let K_n be the complete graph on n vertices, it is easy to see that $\lambda_i(K_n) = n$ for $i \geq 2$. We can also write that

$$L_{K_n} = \sum_{a<b} L_{G_{a,b}}$$

this allows us to prove the following

$$K_n = \sum_{a<b} G_{a,b} \preceq \sum_{a<b} (b-a) P_{a,b} \preceq \sum_{a<b} (b-a) P_n$$
$$\implies \lambda_2(K_n) = n \leq \lambda_2(P_n) \sum_{a<b} (b-a) = \frac{n(n+1)(n-1)}{6} \lambda_2(P_n)$$

Lemma 1.4.4: Bounding λ_2 of the Path Graph

$$\lambda_2(P_n) \geq \frac{6}{(n+1)(n-1)}$$

Let T_d be the complete binary tree of depth d, this will have $n = 2^{d+1} - 1$ vertices. Let $T_d^{a,b}$ be the shortest path between vertices a, b on T_d. Note that this path has size atmost $2d \leq 2\log_2 n$. Doing a similar comparision with K_n we get

$$K_n = \sum_{a<b} G_{a,b} \preceq \sum_{a<b} 2dT_d^{a,b} \preceq \sum_{a<b} 2\log_2 nT_d = \binom{n}{2} 2\log_2 nT_d$$
$$\implies \lambda_2(K_n) = n \leq \binom{n}{2} 2\log_2 n\lambda_2(T_d)$$

Lemma 1.4.5: Bounding λ_2 of Complete Binary Trees

$$\lambda_2(T_d) \geq \frac{1}{(n-1)\log_2 n}$$

2 Random Graphs

An Erdös-Rényi random graph is a graph in which each edge is present with probability p independent of other edges. We can write the adjacency matrix M of this graph as

$$M = p(J - I) + R$$

where J is the all ones matrix, I is the identity matrix and R is defined for off diagonal entries as

$$R(a,b) = \begin{cases} 1-p & \text{with probability } p \\ -p & \text{with probability } 1-p \end{cases}$$

clearly the expectation of M is $\rho(J - I)$, which means the expectation of R is the zero matrix. This can be easily verified from the definition of R. We will show that the eigenvalues of R are usually small and thus M is approximately $p(J - I)$. The following lemma gives bounds on the eigenvalues of M.

Lemma 2.0.1

Let the eigenvalues of R be $\rho_1 \geq ... \geq \rho_n$. Then, the eigenvalues of $R - pI$ are $\rho_i - p$. Since pJ is rank one, the eigenvalues of $M = R - pI + pJ$ interlace the eigenvalues of $R - pI$.

We can get some quantitative results on the eigenvalues of R by calculating the moments of ρ_i. It is easy to verify that the kth moment is

$$\sum_{i=1}^{n} \rho_i^k = Tr(R^k)$$

2.1 Appendix

Theorem 2.1.1

Let A, B be commuting symmetric matrices on a vector space V, then they can be simultaneously diagonalized by the same unitary transformation.

Proof. We will prove by induction on the dimension of V that A, B have the same eigenspaces, orthonormal basis of these eigenspaces will then provide the unitary transformation.

Clearly the statement is true for $\dim V = 1$, let $\dim V = n > 1$ and assume the statement is true for all V such that $\dim V < n$. Let $\alpha_1, ..., \alpha_k$ be the distinct eigenvalues of A. If $k = 1$, then $A = \lambda_1 I$ and thus we are done. So assume $k \geq 2$, let Λ_i be the eigenspace corresponding to the eigenvalue α_i. Then, Λ_i is an invariant subspace of B since for any $\boldsymbol{v} \in \Lambda_i$

$$A(B\boldsymbol{v}) = BA\boldsymbol{v} = \lambda_i B\boldsymbol{v} \implies B\boldsymbol{v} \in \Lambda_i$$

Thus, we can restrict A, B to λ_i and apply the induction hypothesis (since $k \geq 2 \implies \dim \Lambda_i < n$) to get an eigenbasis of Λ_i that simultaneously diagonalizes A, B over Λ_i. The union of these eigenbases over all Λ_i gives us a basis for V that simultaneously diagonalizes both A, B. $\square$

The above theorem is helpful in calculating the eigenvalues of $A + B$.

Theorem 2.1.2: Markov's Inequality

Let X be a non negative random variable with $a > 0$, then

$$P(X > aE(X)) \le a^{-1}$$

Proof.

$$E(X) = P(X \le aE(X))E(X|X \le aE(X)) + P(X > aE(X))E(X|X > aE(X))$$
$$\implies P(X > aE(X)) \le \frac{E(X)}{E(X|X > aE(X))} \le \frac{E(X)}{aE(X)} = a^{-1}$$

□

2.2 Cauchy's Interlacing Theorem for Rank One Updates

The result that we want to prove is

Theorem 2.2.1: Cauchy Interlacing Theorem for Rank One Updates

Let A be a symmetric matrix with eigenvalues $\alpha_1 \ge ... \ge \alpha_n$. Let $B = A + \boldsymbol{x}\boldsymbol{x}^T$ for some vector $\boldsymbol{x}$, let B have eigenvalues $\beta_1 \ge ... \ge \beta_n$, then

$$\beta_i \ge \alpha_i \ge \beta_{i+1}$$

We will do this in two ways. In the first proof, we show that A is a principal submatrix of a matrix with 'similar' eigenvalues to B, this will allow us to apply Cauchy's interlacing theorem for principal submatrices. In the second proof, we will relate the characteristic polynomials and their roots.

2.2.1 The First Proof

We first prove a couple of lemmas that are known as 'Sylvester's theorem'.

Lemma 2.2.2

Let A be an $m \times n$ matrix and B be an $n \times m$ matrix, then we can write

$$x^n \det(xI_m - AB) = x^m \det(xI_n - BA)$$

Proof. Define the following matrices

$$C = \begin{bmatrix} xI_m & A \\ B & I_n \end{bmatrix}, \quad D = \begin{bmatrix} I_m & 0 \\ -B & xI_n \end{bmatrix}$$

we can see that

$$CD = \begin{bmatrix} xI_m - AB & xA \\ 0 & xI_n \end{bmatrix}, \quad DC = \begin{bmatrix} xI_m & A \\ 0 & xI_n - BA \end{bmatrix}$$

using $\det(CD) = \det(DC)$, we get the required result. □

Lemma 2.2.3

Let A be an $m \times n$ matrix and B be an $n \times m$ matrix, then AB and BA have the same set of non zero eigenvalues.

Proof. This follows trivially from the previous lemma or from the following proof. Let $\boldsymbol{x}$ be an eigenvector of AB with eigenvalue $\lambda \neq 0$, it is simple to check that $\boldsymbol{y} = B\boldsymbol{x}$ is an eigenvector of BA with eigenvalue λ. □

Lemma 2.2.4

For positive definite symmetric matrices A, a matrix $\sqrt{A}$ such that $\sqrt{A}^2 = A$ exists.

Proof. Existence follows since the diagonalized matrix of A has a square root. □

We are now ready to prove the theorem.

Proof. We will prove the result for positive definite symmetric matrices A by adding $(1-\alpha_n)I$ to A.
Let $\sqrt{A}$ be a square root of A, now define the $n \times (n+1)$ matrix C as

$$C := \begin{bmatrix} \sqrt{A} & \boldsymbol{x} \end{bmatrix}$$

then we can write

$$CC^T = B, \quad C^TC = \begin{bmatrix} A & \sqrt{A}x \\ x^T\sqrt{A} & x^Tx \end{bmatrix}$$

By previous lemma, C^TC and CC^T share the same set of nonnegative eigenvalues. The eigenvalues of A interlace the eigenvalues of C^TC since it is a principal submatrix. The eigenvalues of C^TC are thus all positive except the smallest (which is 0). Since the eigenvalues of $CC^T = B$ are positive, the result follows. □

2.2.2 The Second Proof

Definition 2.2.1: Cyclic Vectors

Let A be a linear operator on a vector space V. $\boldsymbol{x} \in V$ is called cyclic if the set of finite linear combinations of $\{A^n\boldsymbol{x} \mid n \in \mathbb{N}_0\}$ equals V.

The above definition is equivalent to saying that $\boldsymbol{x} \in V$ is cyclic if

$$V = \{p(A)\boldsymbol{x} \mid p \in \mathbb{R}[t]\}$$

Lemma 2.2.5

A symmetric linear operator A on a vector space V has a cyclic vector iff A has no repeated eigenvalues.

Proof. Suppose that $\boldsymbol{x}$ is a cyclic vector, then

$$V = \{p(A)\boldsymbol{x} \mid p \in \mathbb{R}[t]\}$$

Fix an orthonormal basis on V, since A is symmetric, it is diagonalizable: $A = UDU^T$ where D is diagonal and U is orthogonal. We can then rewrite the above condition as

$$V = \{p(D)\boldsymbol{y} \mid p \in \mathbb{R}[t]\}, \quad (\boldsymbol{y} = U^T\boldsymbol{x})$$

Note that $\{p(D) \mid p \in R[t]\}$ is a vector space with dimension equal to the number of distinct eigenvalues of A. Therefore, $\{p(D)\boldsymbol{y} \mid p \in \mathbb{R}[t]\}$ has the same dimension and we cannot have repeated eigenvalues.

Conversely, suppose A has no repeated eigenvalues, then $\boldsymbol{x} = U\boldsymbol{y}$ corresponding to $\boldsymbol{y} = \mathbf{1}$ is a cyclic vector (it is simple to show that each eigenvector can be written in the form $p(D)\boldsymbol{y}$ for some p, and thus $V = \{p(D)\boldsymbol{y} \mid p \in \mathbb{R}[t]\}$). □

Definition 2.2.2

Let A be a matrix, then $\chi(A)$ represents the characteristic polynomial of A given by

$$\chi(A)(z) = \det(zI - A)$$

Lemma 2.2.6

Let A be a symmetric linear operator on a vector space V, let W be a subspace of V. Then the eigenvalues of A are the same as the union of the eigenvalues of A restricted to W and $W^{\perp}$.

We are now ready to prove the theorem.

Proof. Define $W := \{p(A)\boldsymbol{x} \mid p \in \mathbb{R}[t]\}$ and let $W^{\perp}$ denote the orthogonal vector space. Since $\boldsymbol{x} \in W$, $\boldsymbol{x}^T\boldsymbol{v} = 0$ for all $\boldsymbol{v} \in W^{\perp}$, this means that $A = B$ on $W^{\perp}$. Now, let us retrict A, B to W, let $k = \dim W$. Since W is cyclic by definition, A has k distinct eigenvalues $\lambda_1 > ... > \lambda_k$ with eigenvectors $\boldsymbol{\phi}_i$. Let z not be an eigenvalue of A, then

$$\begin{aligned}\chi(B)(z) &= \det\big(zI - A - \boldsymbol{x}\boldsymbol{x}^T\big)\\ &= \det(zI - A)\det\big(I - (zI - A)^{-1}\boldsymbol{x}\boldsymbol{x}^T\big)\end{aligned}$$

$zI - A$ is invertible since z is not an eigenvalue of A. We can now use Sylvester's theorem with $x = 1$ to write

$$\begin{aligned}\chi(B)(z) &= \chi(A)(z)(1 - \boldsymbol{x}^T(zI - A)^{-1}\boldsymbol{x})\\ &= \chi(A)(z)\left(1 - \sum_{i=1}^{k}\frac{(\boldsymbol{x}^T\boldsymbol{\phi}_i^T)^2}{z - \lambda_i}\right)\end{aligned}$$

Define the function

$$G(z) = \sum_{i=1}^{k}\frac{(\boldsymbol{x}^T\boldsymbol{\phi}_i)^2}{z - \lambda_i}$$

Note the following properties of G

1. Solutions to $G(z) = 1$ are eigenvalues of B.
2. Since $\phi_i \in W$, we must have $\boldsymbol{x}^T\boldsymbol{\phi}_i \neq 0$.
3. There is exactly one solution (μ_{i+1}) to $G(z) = 1$ for $z \in (\lambda_{i+1}, \lambda_i)$, $1 \leq i \leq k - 1$.
4. There is exactly one solution (μ_1) to $G(z) = 1$ for $z > \lambda_1$.

It is now obvious that μ_i are exactly the eigenvalues of B and they interlace the eigenvalues of A.

$$\mu_i \geq \lambda_i \geq \mu_{i+1}$$

Since $A = B$ on $W^{\perp}$, the interlacing property is not affected by including the eigenvalues from this vector space. $\square$