

Lecture 6

Lecturer: David P. Williamson

Scribe: Qingxuan (Max) Jiang

1 More eigenvalue identities

In the first half of this lecture, we will present a few more useful eigenvalue identities.

1.1 Recap of last lecture

We first recall some concepts and properties from last lecture.

General Setting: Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix with real eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ with corresponding eigenvectors x_1, x_2, \dots, x_n such that x_i are orthonormal.

Rayleigh quotient: We have

$$\lambda_k = \min_{x \perp \text{span}(x_1, \dots, x_{k-1})} \frac{x^\top A x}{x^\top x} = \min_{x \in \text{span}(x_k, \dots, x_n)} \frac{x^\top A x}{x^\top x} = \max_{x \in \text{span}(x_1, \dots, x_k)} \frac{x^\top A x}{x^\top x}$$

Representaion of matrix and its inverse via eigenvalues and eigenvectors:

$$\begin{aligned} A &= \lambda_1 x_1 x_1^\top + \lambda_2 x_2 x_2^\top + \dots + \lambda_n x_n x_n^\top \\ I &= x_1 x_1^\top + x_2 x_2^\top + \dots + x_n x_n^\top \\ A^{-1} &= \frac{1}{\lambda_1} x_1 x_1^\top + \frac{1}{\lambda_2} x_2 x_2^\top + \dots + \frac{1}{\lambda_n} x_n x_n^\top \end{aligned}$$

Diagonalization of a matrix: Let $X = \begin{bmatrix} | & \dots & | \\ x_1 & \dots & x_n \\ | & \dots & | \end{bmatrix}$ and $D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$.

Since x_i are orthonormal to each other, we have $X^\top X = I$, $X^\top = X^{-1}$. Moreover, from definition of eigenvalues we have $AX = XD$, so multiplying by $X^{-1} = X^\top$ on the right gives $A = XDX^{-1} = XDX^\top$.

1.2 More simple eigenvalue properties

We first investigate properties of eigenvalues and eigenvectors for matrix powers A^k , where A^k is A multiplied by itself k times.

Lemma 1 *The eigenvectors of A^k are x_1, \dots, x_n with corresponding eigenvalues $\lambda_1^k, \dots, \lambda_n^k$.*

Proof: From the observations above, we can write $A^k = (XDX^{-1})^k = XD^kX^{-1}$. Thus, we have $A^kX = XD^k$. By analyzing corresponding columns of both sides of the equation, we can see that $A^kx_i = \lambda_i^kx_i$ for any $i \in \{1, \dots, n\}$, and thus x_1, \dots, x_n are eigenvalues of A^k with eigenvalues $\lambda_1^k, \dots, \lambda_n^k$. \square

For our next few identities, we will assume the following fact without proof:

Fact 1 $\det(AB) = \det(A)\det(B)$

A simple corollary of this fact is as follows:

Corollary 2 $\det(A^{-1}) = \frac{\det(A)\det(A^{-1})}{\det(A)} = \frac{\det(AA^{-1})}{\det(A)} = \frac{\det(I)}{\det(A)} = \frac{1}{\det(A)}$

We now derive eigenvalue representations of the determinant:

Lemma 3 $\det(A) = \prod_{i=1}^n \lambda_i$

Proof:

$$\begin{aligned} \det(A) &= \det(XDX^{-1}) = \det(X)\det(D)\det(X^{-1}) \\ &= \det(D)\det(XX^{-1}) = \det(D) = \prod_{i=1}^n \lambda_i. \end{aligned}$$

\square

Recall that in the first lecture, we defined the *trace* of A to be $\text{tr}(A) = \sum_{i=1}^n a_{ii}$.

Lemma 4 $\text{tr}(A) = \sum_{i=1}^n \lambda_i$.

Proof: Consider the characteristic polynomial of A , which we defined in the first lecture to be $\det(\lambda I - A)$. This is a degree n polynomial in λ . The idea of the proof is to consider the coefficient of λ^{n-1} and represent it in two different ways using a_{ii} and λ_i , and then equating these two representations.

We first write $\det(\lambda I - A)$ in terms of eigenvalues λ_i :

$$\begin{aligned} \det(\lambda I - A) &= \det(\lambda XX^T - XDX^T) = \det(X(\lambda I - D)X^T) \\ &= \det(X)\det(\lambda I - D)\det(X^T) = \det(\lambda I - D) = \prod_{i=1}^n (\lambda - \lambda_i). \end{aligned}$$

In this case, the coefficient of λ^{n-1} is $-\sum_{i=1}^n \lambda_i$.

We then write $\det(\lambda I - A)$ in terms of the entries a_{ii} . Recall that we can represent the determinant of a matrix Z as a sum of multiplied entries with indices in the set of

permutations S_n : $\det(Z) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{i=1}^n a_{i\sigma(i)}$. We then look at our determinant:

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}.$$

We see that only permutation that can produce a λ^{n-1} term in the above sum is the identity $(\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn})$. (Every other permutation would choose at least two terms not on the diagonal, thus have highest power λ^{n-2} or lower.)

Thus, we know that the coefficient of λ^{n-1} term can also be represented as $-\sum_{i=1}^n a_{ii}$.

Equating the two cases above, we know that

$$\text{tr}(A) = \sum_{i=1}^n a_{ii} = \sum_{i=1}^n \lambda_i.$$

□

1.3 Eigenvalue Interlacing Theorem

With the above lemmas, we are able to prove the eigenvalue interlacing theorem, which we have already used without proof in Lecture 2 on proving Wilf's theorem on chromatic numbers and Huang's theorem on the sensitivity conjecture.

Theorem 5 (Eigenvalue Interlacing Theorem) *Suppose $A \in \mathbb{R}^{n \times n}$ is symmetric. Let $B \in \mathbb{R}^{m \times m}$ with $m < n$ be a principal submatrix (obtained by deleting both i -th row and i -th column for some values of i). Suppose A has eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$ and B has eigenvalues $\beta_1 \leq \cdots \leq \beta_m$. Then*

$$\lambda_k \leq \beta_k \leq \lambda_{k+n-m} \quad \text{for } k = 1, \dots, m$$

And if $m = n - 1$,

$$\lambda_1 \leq \beta_1 \leq \lambda_2 \leq \beta_2 \leq \cdots \leq \beta_{n-1} \leq \lambda_n$$

Proof: WLOG, we can assume $A = \begin{bmatrix} B & X^\top \\ X & Z \end{bmatrix}$. To see why this is WLOG, suppose we switch the i -th and j -th row and the i -th and j -th column of A at the same time to obtain A' . Then A' will have the same eigenvalues as A , with corresponding eigenvectors also switching the i -th and j -th entry. Thus, if the principal submatrix is not in the first m rows/columns, we can exchange its row/columns with the first m rows/columns without changing the eigenvalues.)

Let $\{x_1, \dots, x_n\}$ be orthonormal eigenvectors of A , and $\{y_1, \dots, y_m\}$ be orthonormal eigenvectors of B .

We first show $\lambda_k \leq \beta_k$. We define the following vector spaces:

$$V = \text{span}(x_k, \dots, x_n), \quad W = \text{span}(y_1, \dots, y_k), \quad \widetilde{W} = \left\{ \begin{bmatrix} w \\ 0 \end{bmatrix} \in \mathbb{R}^n, w \in W \right\}$$

Since $\dim(V) = n - k + 1$ and $\dim(\widetilde{W}) = \dim(W) = k$, there exists $\tilde{w} \in V \cap \widetilde{W}$ and $\tilde{w} = \begin{bmatrix} w \\ 0 \end{bmatrix}$ for some $w \in W$. Then

$$\tilde{w}^T A \tilde{w} = [w^T \ 0] \begin{bmatrix} B & X^T \\ X & Z \end{bmatrix} \begin{bmatrix} w \\ 0 \end{bmatrix} = w^T B w$$

Recall $\lambda_k = \min_{x \in V} \frac{x^T A x}{x^T x}$ and $\beta_k = \max_{x \in W} \frac{x^T B x}{x^T x}$. Then we see that

$$\lambda_k \leq \frac{\tilde{w}^T A \tilde{w}}{\tilde{w}^T \tilde{w}} = \frac{w^T B w}{w^T w} \leq \beta_k.$$

We then show $\beta_k \leq \lambda_{k+n-m}$. Similarly, we define the vector spaces

$$V = \text{span}(x_1, \dots, x_{k+n-m}), \quad W = \text{span}(y_k, \dots, y_m), \quad \widetilde{W} = \left\{ \begin{bmatrix} w \\ 0 \end{bmatrix} \in \mathbb{R}^n, w \in W \right\}$$

Since $\dim(V) = k + n - m$, $\dim(\widetilde{W}) = \dim(W) = m - k + 1$, there exists $\tilde{w} \in V \cap \widetilde{W}$ and $\tilde{w} = \begin{bmatrix} w \\ 0 \end{bmatrix}$ for some $w \in W$. As before, we have $\tilde{w}^T A \tilde{w} = w^T B w$. It follows that

$$\lambda_{k+n-m} = \max_{x \in V} \frac{x^T A x}{x^T x} \geq \frac{\tilde{w}^T A \tilde{w}}{\tilde{w}^T \tilde{w}} = \frac{w^T B w}{w^T w} \geq \min_{x \in W} \frac{x^T B x}{x^T x} = \beta_k,$$

completing the proof. \square

2 Bipartite Graphs

In the second half of the lecture, we will show how all the various eigenvalue identities we've proven over this lecture and the last can be applied to showing something about the structure of graphs. In particular, we show that the spectrum of the adjacency matrix tells us whether the graph is bipartite or not.

Lemma 6 *If G is bipartite, and λ is an eigenvalue of adjacency matrix A , then so is $-\lambda$.*

Proof: If G is bipartite, we can re-index the nodes such that

$$A = \begin{bmatrix} 0 & B \\ B^T & 0 \end{bmatrix}.$$

Let $v = \begin{bmatrix} x \\ y \end{bmatrix}$ be an eigenvector of A with eigenvalue λ . Then we have

$$A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}$$

Hence we have $By = \lambda x$ and $B^T x = \lambda y$. From this, we have

$$\begin{bmatrix} 0 & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} x \\ -y \end{bmatrix} = \begin{bmatrix} -By \\ B^T x \end{bmatrix} = \begin{bmatrix} -\lambda x \\ \lambda y \end{bmatrix} = -\lambda \begin{bmatrix} x \\ -y \end{bmatrix}$$

So, $-\lambda$ is an eigenvalue corresponding to the eigenvector $\begin{bmatrix} x \\ -y \end{bmatrix}$. □

We can now show that this statement can be made an “if and only if”: that is the graph G is bipartite if and only if for each eigenvalue λ there is another eigenvalue $-\lambda$ with multiplicity; that is, if the multiplicity of λ is k , then so is the multiplicity of $-\lambda$.

Theorem 7 *If for each eigenvalue $\lambda \neq 0$ there is another eigenvalue $\lambda' = -\lambda$ (with multiplicity), then G is bipartite.*

Proof: Let k be any odd positive integer. Recall from Lemma 1 that if A has eigenvalues $\lambda_1, \dots, \lambda_n$, then A^k has eigenvalues $\lambda_1^k, \dots, \lambda_n^k$. From our assumption, we have

$$\text{tr}(A^k) = \sum_{i=1}^n \lambda_i^k = 0$$

Also recall that $(A^k)_{ij}$ is the number of walks from i to j of length exactly k . So if there is an odd cycle of length k , then it must be the case that $(A^k)_{ii} > 0$, so that $\text{tr}(A^k) > 0$, a contradiction to $\text{tr}(A^k) = 0$. Thus, there are no odd cycles of length k . Since this is true for any odd positive integer k , there are no odd cycles in G , which implies that G is bipartite. □

Now we can show something even stronger than the previous statement: we only need to look at the smallest and largest eigenvalue to know whether or not the graph is bipartite.

Theorem 8 *Suppose G is connected. Then, $\lambda_1 = -\lambda_n$ if and only if G is bipartite.*

Proof: We have already seen in Lemma 6 that if G is bipartite, then A must have $\lambda_n = -\lambda_1$ (as they must form a pair of eigenvalues with largest absolute value). We now assume that $\lambda_1 = -\lambda_n$. Let x_1 be the eigenvector corresponding to λ_1 with $x_1^T x_1 = 1$. Let $z \in \mathbb{R}^n$ with $z(i) = |x_1(i)|$, then we have the following:

$$|\lambda_1| = |x_1^T A x_1| \leq \sum_{i,j} a_{ij} |x_1(i)| |x_1(j)| = \sum_{i,j} a_{ij} z(i) z(j) = z^T A z \leq \max_x \frac{x^T A x}{x^T x} = \lambda_n.$$

Since $\lambda_1 = -\lambda_n$, equality holds for all inequalities in the above expression, so (a) it must be that z is the eigenvector associated with λ_n ; and (b) $\forall i, j, a_{ij} x_1(i) x_1(j) \leq 0$, so $x_1(i) x_1(j) \leq 0$ for every edge (i, j) .

If we can assume that $z > 0$, then from (b) we know that every edge (i, j) has one of $x_1(i), x_1(j)$ positive, the other negative. This implies the following partition

$$V = \{i : x_1(i) < 0\}, \quad W = \{i : x_1(i) > 0\},$$

which shows that G is a bipartite graph.

We finally go back and show that $z > 0$. We know that $z \geq 0$, and we assume that there exists some entry of z that is 0. Since G is connected and $z \neq 0$, there exists some edge (i, k) such that $z(i) = 0$ and $z(k) \neq 0$. (If no such edge exists, then $\{i : z(i) = 0\}$ and $\{k : z(k) \neq 0\}$ would be two nonempty, disconnected components of G , a contradiction.) Then we have $(Az)(i) = \sum_{j:(i,j) \in E} z(j) > 0$. But from (a) we also have $(Az)(i) = \lambda_n z(i) = 0$, and this gives a contradiction. Thus, we must have $z > 0$. This completes our proof. \square