10 Eigenvalue inequalities; graph laplacians

Wednesday, October 7, 2020 11:04 PM

Prop 16.14/16.26 (Poincare separation thm) (applications to quantum mechanics)

Let A be a $n \times n$ symmetric (or Hermitian) matrix, let $r \in \mathbb{Z}$ with $1 \le r \le n$, and let $(u_1, \dots, u_r)$ be r orthonormal vectors.

Let $B = (u_i^T A u_j)$ (an $r \times r$ matrix),

let $\lambda_1(A) \le \dots \le \lambda_n(A)$ be eigenvalues of A

$\lambda_1(B) \le \dots \le \lambda_r(B)$ be eigenvalues of B

Then $\lambda_k(A) \le \lambda_k(B) \le \lambda_{k+n-r}(A)$, $k = 1, \dots, r$.

proof. Immediate corollary of prev. Prop. (eigenvalue interlacing) letting $R = (u_1 \cdots u_r)$. $\boxed{/}$

We also have an immediate corollary of interlacing eigenvalues of matrix minors.

Let $P_1 \in \mathbb{R}^{n \times n-1}$ be defined by $I_{n\times n}[1:n; 1:n-1]$, the identity minus the last col.

Then $P_1^T P_1 = I$ and $B = P_1^T A P$ is $A[1:n-1; 1:n-1]$.

Then we get a genuine interlacing

$$\lambda_1 \le \mu_1 \le \lambda_2 \le \mu_2 \le \dots \le \mu_{n-2} \le \lambda_{n-1} \le \mu_{n-1} \le \lambda_n .$$

We get of course similar results for $B = A[1:r; 1:r]$,

$$\lambda_k \le u_k \le \lambda_{k+n-r}, \quad k = 1, \dots, r.$$

(Cauchy interlacing theorem)

Thm 16.13/16.27 (Courant-Fischer) Let A be a symmetric $n \times n$ matrix with eigenvalues $\lambda_1 \le \dots \le \lambda_n$. If $\mathcal{V}_k$ denotes the set of subspaces of $\mathbb{R}^n$ of dim k, then

$$\lambda_k = \max_{W \in \mathcal{V} \dots} \min_{x \in W, x \neq 0} \frac{x^T A x}{x^T x} .$$ (min-max thm)

$$\lambda_k = \max_{W \in \mathcal{V}_{n-k+1}} \min_{x \in W,\, x \neq 0} \frac{x^T A x}{x^T x}.$$ (min-max thm)

$$\lambda_k = \min_{W \in \mathcal{V}_k} \max_{x \in W,\, x \neq 0} \frac{x^T A x}{x^T x}.$$

proof. The proofs for both equations are similar, so we'll just prove the second.

Let $(u_1, \dots, u_n)$ be an orthonormal basis of eigenvectors corresponding to $\lambda_1, \dots, \lambda_n$.

Let $V_k = \operatorname{span}\{u_1, \dots, u_k\}$, so $\dim(V_k) = k$.

By Rayleigh-Ritz,

$$\lambda_k = \max_{\substack{x \neq 0 \\ x \in V_k}} \frac{x^T A x}{x^T x} \geq \min_{W \in \mathcal{V}_k} \max_{\substack{x \neq 0 \\ x \in W}} \frac{x^T A x}{x^T x},$$ proving the easy direction.

Lemma: For any $W \in \mathcal{V}_k$, $\dim(W \cap V_{k-1}^{\perp}) \geq 1$.

proof. $\dim(V_{k-1}) = k-1$, so $\dim(V_{k-1}^{\perp}) = n-k+1$

But recall that

$$\dim(A) + \dim(B) = \dim(A \cap B) + \dim(A+B),$$ for A, B subspaces of E.

Thus,

$$\dim(W) + \dim(V_{k-1}^{\perp}) = \dim(W \cap V_{k-1}^{\perp}) + \underbrace{\dim(W + V_{k-1}^{\perp})}_{\leq \mathbb{R}^n}$$

$$\Rightarrow \quad k + n - k + 1 \leq \dim(W \cap V_{k-1}^{\perp}) + n$$

$$\Rightarrow \quad \dim(W \cap V_{k-1}^{\perp}) \geq 1. \qquad \square$$

Thus, $\exists\, v \in W \cap V_{k-1}^{\perp}$ s.t. $v \neq 0$.

By Rayleigh-Ritz,

$$\lambda_k = \min_{\substack{x \neq 0 \\ x \in V_{k-1}^{\perp}}} \frac{x^T A x}{x^T x} \leq \frac{v^T A v}{v^T v} \leq \max_{\substack{x \in W \\ x \neq 0}} \frac{x^T A x}{x^T x}, \quad \forall\, W \in \mathcal{V}_k.$$

$$\Rightarrow \quad \lambda_k \leq \min_{W \in \mathcal{V}_k} \max_{x \in W} \frac{x^T A x}{x^T x}. \qquad \square$$

$$\Rightarrow \quad \lambda_k \le \min_{W \in \mathcal{V}_k} \max_{\substack{x \in W \\ x \ne 0}} \frac{x^T A x}{x^T x}.$$

Courant-Fischer is useful for proving results about perturbations of eigenvalues

Prop. 16.15/16.28 Given symmetric $A, B \in \mathbb{R}^{n \times n}$, $B = A + \Delta A$, if

$\alpha_1 \le \dots \le \alpha_n$ are eigenvalues of A

$\beta_1 \le \dots \le \beta_n$ are eigenvalues of B,

then

$$|\alpha_k - \beta_k| \le \rho(\Delta A) \le \|\Delta A\|_2 .$$

Proof. We use Courant-Fischer applied to B.

$$\beta_k = \min_{W \in \mathcal{V}_k} \max_{\substack{x \in W \\ x \ne 0}} \frac{x^T B x}{x^T x}$$

$$\le \max_{\substack{x \in V_k \\ x \ne 0}} \frac{x^T B x}{x^T x} = \max_{\substack{x \in V_k \\ x \ne 0}} \left(\frac{x^T A x}{x^T x} + \frac{x^T \Delta A x}{x^T x} \right)$$

$$\le \max_{x \in V_k} \frac{x^T A x}{x^T x} + \max_{x \in V_k} \frac{x^T \Delta A x}{x^T x}$$

$$= \alpha_k + \max_{x \in V_k} \frac{x^T \Delta A x}{x^T x}$$

$$\le \alpha_k + \max_{x \in \mathbb{R}^n} \frac{x^T \Delta A x}{x^T x} .$$

$$\le \alpha_k + \rho(\Delta A)$$

$$\Rightarrow \quad \beta_k \le \alpha_k + \rho(\Delta A)$$

Swapping A and B, $\quad \alpha_k \le \beta_k + \rho(\Delta A)$

$$\Rightarrow \quad |\alpha_k - \beta_k| \le \rho(\Delta A) \le \|\Delta A\|_2, \quad k = 1, \dots, n.$$

Can get other similar results, like Weyl's inequalities.

Prop. ~~16.6~~/16.29 (Weyl) Given two symmetric (or Hermitian) $n \times n$ matrices A and B, the following inequalities hold. For all i, j, k with $1 \le i, j, k \le n$:

(1) If $i + j = k + 1$, then $\lambda_i(A) + \lambda_j(B) \le \lambda_k(A+B)$

(2) If $i + j = k + n$, then $\lambda_k(A+B) \le \lambda_i(A) + \lambda_j(B)$.

proof. Note that if we let $A \leftarrow -A$ and $B \leftarrow -B$, then (2) $\rightarrow$ (1).

By Courant-Fischer, $\exists$ subspace H with $\dim(H) = n-k+1$ s.t.

$$\lambda_k(A+B) = \min_{\substack{x \in H \\ x \neq 0}} \frac{x^T(A+B)x}{x^T x}.$$

Similarly, $\exists$ subspaces F and G, $\dim(F) = i$, $\dim(G) = j$ s.t.

$$\lambda_i(A) = \max_{\substack{x \in F \\ x \neq 0}} \frac{x^T A x}{x^T x}, \qquad \lambda_j(B) = \max_{\substack{x \in G \\ x \neq 0}} \frac{x^T A x}{x^T x}.$$

Lemma: $\dim(F \cap G \cap H) \ge 1$.

proof. Use Grassman relation twice. (like in prev. lemma) $\square$

$\Rightarrow$ $\exists$ unit vector $z \in F \cap G \cap H$, so

$$\lambda_k(A+B) \le z^T(A+B)z, \quad \lambda_i(A) \ge z^T A z, \quad \lambda_j(B) \ge z^T B z.$$

$$\Rightarrow \lambda_k(A+B) \le \lambda_i(A) + \lambda_j(B).$$

$\square$

Corollaries: $\lambda_1(A) + \lambda_k(B) \le \lambda_k(A+B)$

$\lambda_k(A+B) \le \lambda_k(A) + \lambda_n(B)$

$$\Rightarrow \lambda_1(A) + \lambda_k(B) \le \lambda_k(A+B) \le \lambda_k(A) + \lambda_n(B)$$

This implies the monotonicity theorem for symmetric (and Hermitian) matrices.

If A is symmetric and B is symm pos semidef, then symmetric (and Hermitian) matrices.

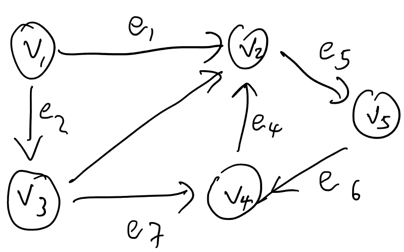
If A is symmetric, and B is symm. pos. semidef., then

$$\lambda_k(A) \le \lambda_k(A+B).$$

Graphs and graph laplacians

Def 18.1 A *directed graph* is a pair $G=(V,E)$, where $V=\{v_1,\dots,v_m\}$ is a set of *nodes/vertices*, $E \subseteq V \times V$ is a set of ordered pairs (u,v), with $u \ne v$, called *edges*. Given any edge $e=(u,v)$, let $s(e)=u$ be the *source* of e and $t(e)=v$ be the *target* of E.

Ex $G_1 =$



Ex $D(G_1) = \begin{bmatrix} 2 & & & & \\ & 4 & & 0 & \\ & & 3 & & \\ & 0 & & 3 & \\ & & & & 2 \end{bmatrix}$

Def. 18.2 For every node $v \in V$, the degree $d(v)$ is the number of edges leaving or entering v:

$$d(v) = \left| \{u \in V \mid (v,u) \in E \text{ or } (u,v) \in E\} \right|.$$

Notation: $d(v_i) = d_i$.

The degree matrix $D(G) = \operatorname{diag}(d_1,\dots,d_m)$

Def. 18.3 Given a directed graph $G=(V,E)$, a *path* from u to v $(u,v \in V)$ is a sequence of nodes $(v_0,\dots,v_k)$ s.t. $u=v_0$, $v=v_k$ and $(v_i, v_{i+1}) \in E$ $\forall\, 0 \le i \le k-1$. k is the *length* of the path.

A path is *closed* if $u=v$. G is *strongly connected* if $\forall u \ne v \in V$, there is a path $u \to v$ and a path $v \to u$.

If $\exists$ paths $u \to v$ and $v \to u$, then u and v are *related*.

F... $\in V$ the *strongly connected component* containing v

If $\exists$ paths $u \to v$ and $v \to u$, then u and v are ...

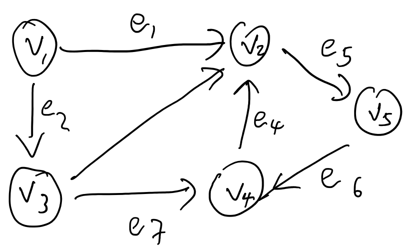
For any $v \in V$, the strongly connected component containing v is the set of all nodes related to v.

Def. 18.4 Given a directed graph $G=(V, E)$, $V=\{v_1, \dots, v_m\}$, $E=\{e_1, \dots, e_n\}$, the incidence matrix $B(G)$ of G is the $m \times n$ matrix whose entries

$$b_{ij} = \begin{cases} +1 & \text{if } s(e_j)=v_i \\ -1 & \text{if } t(e_j)=v_i \\ 0 & \text{otherwise} \end{cases}$$

(sometimes just say $B = B(G)$)

Ex $G_1 =$

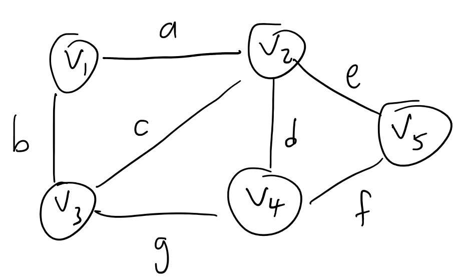


$$B(G_1) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 \end{pmatrix}$$

We can interpret $B : \mathbb{R}^E \to \mathbb{R}^V$ as the boundary map taking edges to vertices (connections to the boundary operator of simplicial complexes and algebraic topology).

Def. 18.5 A graph (or undirected graph) is a pair $G=(V, E)$, where $V = \{v_1, \dots, v_m\}$ is a set of nodes/vertices and E is a set of two-element subsets of V called edges.

Ex. $G_2 =$



$$D(G_2) = \begin{pmatrix} 2 & & & & \\ & 4 & & 0 & \\ & & 3 & & \\ & 0 & & 3 & \\ & & & & 2 \end{pmatrix}$$

Def. 18.6 The degree $d(v) = |\{u \in V \mid \{u,v\} \in E\}|$.

Def. 18.7 Given a (undirected) graph $G=(V, E)$, for any two nodes $u, v \in V$, a path from $u \to v$ is a sequence of nodes $(v_0, v_1, \dots, v_k)$ s.t.

Def. 18.7 Given a (undirected) graph $G=(V,E)$, for any two nodes $u, v \in V$, a *path* from $u \to v$ is a sequence of nodes $(v_0, v_1, \dots, v_k)$ s.t. $u = v_0$, $v_k = v$ and $\{v_i, v_{i+1}\} \in E$ for $0 \le i \le k-1$. k is the *length*.

A path is *closed* if $u = v$. G is *connected* if $\forall u \neq v$, $\exists$ a path $u \to v$.

Connected components are equivalence classes of *related* nodes, where nodes are related if there is a path b/t them.

Def 18.8 The incidence matrix $B = B(G) \in \mathbb{R}^{m \times n}$ is

$$b_{ij} = \begin{cases} +1 & \text{if } e_j = \{v_i, v_k\} \text{ for some } k \\ 0 & \text{otherwise.} \end{cases}$$

Def. 18.9 If $G=(V,E)$ is a directed or an undirected graph, if there is an edge (u,v) in the directed case or $\{u,v\}$ in the undirected case, then v is *adjacent* to u, denoted

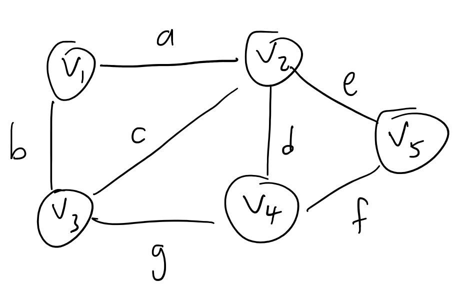
$$u \sim v.$$

Note: $\sim$ is symmetric if G is undirected, but not in general if G is directed.

Def. 18.10 Given a directed or undirected graph $G=(V,E)$, with $V = \{v_1, \dots, v_m\}$, the *adjacency matrix* $A = A(G)$ is the symmetric matrix with

$$a_{ij} = \begin{cases} 1 & \text{if } u \sim v \text{ or } v \sim u \\ 0 & \text{otherwise.} \end{cases}$$

Ex.

$G_2 =$ 

$$A(G_2) = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

We can view $A : \mathbb{R}^V \to \mathbb{R}^V$ s.t. $\forall x \in \mathbb{R}^m$, $(Ax)_i = \sum_{j \sim i} x_j$.

This is a *diffusion operator*, as the value of Ax at v_i is the sum of the values of x at nodes v_j adjacent to v_i.

This gives a geometric interpretation of eigenvalues of A, because if $Ax = \lambda x$, then $\lambda x_i = \sum_{j \sim i} x_j$, so the sum of the values of x at nodes adjacent to i must be equal to λ times the value at v_i.

Def. 18.1 Given any undirected graph $G = (V, E)$, an *orientation* of G is a function $\sigma : E \to V \times V$ assigning a source and target for each edge in E. i.e. $\forall \{u,v\} \in E$, either $\sigma(\{u,v\}) = (u,v)$ or $\sigma(\{u,v\}) = (v,u)$. The *oriented graph* G^σ is the directed graph $G^\sigma = (V, E^\sigma)$, $E^\sigma = \sigma(E)$.

Prop. 18.1 Let $G = (V, E)$ be an undirected graph with $|V| = m$, $|E| = n$, and c connected components. $\forall$ orientation σ of G, if B is the incidence matrix of G^σ, then $c = \dim(\mathrm{Ker}(B^T))$, and $\mathrm{rank}(B) = m - c$. Furthermore, $\mathrm{Ker}(B^T)$ has a basis consisting of indicator vectors of the connected components of G; i.e. vectors $(z_1, \dots, z_m)$ s.t. $z_j = 1$ iff v_j is in the ith component K_i of G and $z_j = 0$ otherwise.

proof. Let $z \in \mathrm{Ker}\, B^T$, so $B^T z = 0 \iff z^T B = 0$.

$\forall$ edge $\{v_i, v_j\}$ of G, the corresponding col of B has 0 entries except for ± 1 at positions i and j.

$\Rightarrow z_i = z_j$.

Thus, if $i \sim j$, $z_i = z_j$, so z has constant value on each connected component of G.

$\Rightarrow$ ... $\{1$ iff $v_j \in K_i$ (ith connected ...)

on each connected component of G.

$\Rightarrow z = \lambda_1 z^1 + \dots + \lambda_c z^c$, where $z^i_j = \begin{cases} 1 & \text{iff } v_j \in K_i \\ 0 & \text{otherwise} \end{cases}$ (ith connected component)

$\Rightarrow \dim(\operatorname{Ker} B^T) = c$ and $\operatorname{Ker} B^T$ has a basis of indicator vectors.

But $m = \dim(\operatorname{Ker}(B^T)) + \operatorname{rank}(B^T)$, so $\operatorname{rank}(B^T) = m - c$.

$\Rightarrow \operatorname{rank}(B) = m - c$. $\boxtimes$

Aside: $B^T \vec{1} = \vec{0}$ always, so $\operatorname{Ker} B^T \neq (0)$.

Def. 18.13 A *weighted graph* is a pair $G = (V, W)$, where $V = \{v_1, \dots, v_m\}$ is a set of *nodes/vertices* and W is a symmetric matrix called the *weight matrix* s.t. $w_{ij} \geq 0$ and $w_{ii} = 0$. A set $\{v_i, v_j\}$ is an *edge* iff $w_{ij} > 0$. The *underlying graph* of G is

(V, E), $E = \{ \{v_i, v_j\} \mid w_{ij} > 0 \}$.

Generalization of undirected graphs with $d(v_i) = \sum_{j=1}^{m} w_{ij}$ degree.

Note $W\vec{1} = \begin{pmatrix} d_1 \\ \vdots \\ d_m \end{pmatrix}$.

Prop. 18.2 Given any undirected graph G, $\forall$ orientation σ of G, if A is the adjacency matrix of G^σ and D is the diagonal degree matrix, then

$$L = B^\sigma (B^\sigma)^T = D - A.$$

proof. Let $B^\sigma = \begin{bmatrix} b_1^\sigma \\ \vdots \\ b_m^\sigma \end{bmatrix}$, b_j^σ is the jth row of B^σ.

Then $\left(B^\sigma (B^\sigma)^T\right)_{ij} = \langle b_i^\sigma, b_j^\sigma \rangle$.

Note $\langle b_i^\sigma, b_i^\sigma \rangle = d(v_i)$ because we get $\langle b_i^\sigma, b_i^\sigma \rangle_k = \sum (b_{ik}^\sigma)^2$.

And if $i \neq j$, $\langle b_i^\sigma, b_j^\sigma \rangle \neq 0$ iff $\exists$ an edge $e_k = (v_i, v_j)$ or $e_k = (v_j, v_i)$.

because if $b_{ik}^\sigma = 1$, then $b_{jk}^\sigma = -1$ or 0

Note that there can only be 1 such undirected edge, so if $\langle b_i^\sigma, b_j^\sigma \rangle \neq 0$, then $\langle b_i^\sigma, b_j^\sigma \rangle = -1$.

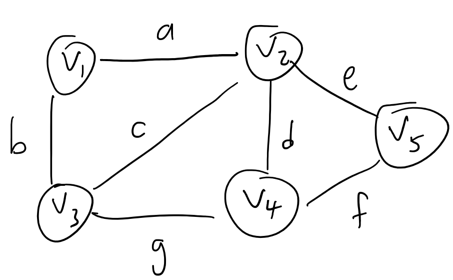
Thus, $L = B^\sigma (B^\sigma)^T = D - A$.

Clearly, L is positive semidefinite as $L = B^\sigma (B^\sigma)^T$, and thus the eigenvalues of L are ≥ 0.

Def. 18.15 The matrix $L = B^\sigma (B^\sigma)^T = D - A$ is called the (unnormalized) graph Laplacian of the graph G^σ and $L = D - A$ is the (unnormalized) graph Laplacian of G.

Ex.

$G_2 =$



$$A(G_2) = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

$$L = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 4 & -1 & -1 & -1 \\ -1 & -1 & 3 & -1 & 0 \\ 0 & -1 & -1 & 3 & -1 \\ 0 & -1 & 0 & -1 & 2 \end{bmatrix}.$$

Def. 18.16 Given a weighted graph $G=(V, W)$, the unnormalized graph Laplacian $L = L(G) = D(G) - W$, where $D(G)$ is the diagonal degree matrix.

There are analogues of incidence matrices and the various other properties to weighted graphs as well.