[CS1961: Lecture 16] Cauchy Interlacing Theorem, Huang's Proof of Sensitivity Conjecture

1 Cauchy Interlacing Theorem

Theorem 1 (Cauchy Interlacing Theorem)**.** *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Let $B \in \mathbb{R}^{m\times m}$ be a principal submatrix*[1] *of A with eigenvalues $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_m$. Then for all $k \in m$, $\lambda_k \geq \mu_k \geq \lambda_{k+n-m}$.*

[1] A principal submatrix is a square submatrix obtained by removing certain rows and columns. The indices of removed rows are the same with removed columns.

Proof. It is sufficient to prove the case when $m = n - 1$. W.l.o.g, assume B is generated by deleting the first row and first column in A. By the Courant-Ficher theorem, we have

$$\lambda_k = \max_{\substack{V\subseteq\mathbb{R}^n \\ \dim(V)=k}} \min_{\mathbf{x}\in V\setminus\{\mathbf{0}\}} R_A(\mathbf{x}) \quad \text{and} \quad \mu_k = \max_{\substack{U\subseteq\mathbb{R}^{n-1} \\ \dim(U)=k}} \min_{\mathbf{y}\in U\setminus\{\mathbf{0}\}} R_B(\mathbf{y}).$$

For any $\mathbf{y} \in \mathbb{R}^{n-1}$, let $\mathbf{y}' = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix}$. Then $\mathbf{y}' \in \mathbb{R}^n$ and $\mathbf{y}'$ satisfies that $\langle \mathbf{y}, B\mathbf{y}\rangle = \langle \mathbf{y}', A\mathbf{y}'\rangle$ and $\langle \mathbf{y}, \mathbf{y}\rangle = \langle \mathbf{y}', \mathbf{y}'\rangle$. Therefore,

$$\mu_k = \max_{\substack{U\subseteq\mathbb{R}^{n-1} \\ \dim(U)=k}} \min_{\mathbf{y}\in U\setminus\{\mathbf{0}\}} R_A(\mathbf{y}').$$

This indicates that $\mu_k \leq \lambda_k$.

For the proof of $\mu_k \geq \lambda_{k+n-m}$, consider $-A$ and $-B$. Then the spectrum of $-A$ is $-\lambda_n \geq -\lambda_{n-1} \geq \cdots \geq -\lambda_1$ and the spectrum of $-B$ is $-\mu_{n-1} \geq -\mu_{n-2} \geq \cdots \geq -\mu_1$. With the same argument, we can verify that $-\mu_k \leq -\lambda_{k+1}$.

□

Let $n^+(A)$ be the number of positive eigenvalues of A and $n^-(A)$ be the number of negative eigenvalues. Let $\alpha(G)$ be the independent number of G. We can derive an upper bound of $\alpha(G)$ using the Cauchy interlacing theorem.

Theorem 2 (Cvetkovic Theorem)**.** $\alpha(G) \leq \min\{n - n^+(A), n - n^-(A)\}$.

Proof. Let $S \subseteq V$ be an independent set with size $\alpha(G)$. Let B be the principal submatrix of A indexed by S. Then B must be a zero matrix with each eigenvalue $\mu_k = 0$ for $k \in [\alpha(G)]$.

For $k \in [\alpha(G)]$, by the Cauchy interlacing theorem, $\lambda_k \geq \mu_k \geq \lambda_{k+n-\alpha(G)}$. Therefore we have

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{\alpha(G)} \geq 0.$$

This indicates $n^-(A) \leq n - \alpha(G)$. Apply the same argument on $-A$ and $-B$, we can similarly yield that $n^+(A) \leq n - \alpha(G)$.

□

Let $\chi(G)$ be the chromatic number of G. We can also derive an upper bound for $\chi(G)$ using Cauchy interlacing theorem.

Theorem 3 (Wilf Theorem). $\chi(G) \leq \lfloor \mu_1 \rfloor + 1$.

When the graph G is d-regular, the Wilf theorem indicates that $\chi(G) \leq d+1$, which is tight in this case.

Proof. We prove this by induction on n. When $n = 1$, the bound obviously holds. When $n \geq 2$, choose a vertex v with the smallest degree. Therefore,

$$\deg(v) \leq d_{\text{ave}} \leq \mu_1(G).$$

Let H be the graph induced by $V \setminus \{v\}$. By induction hypothesis, we have $\chi(H) \leq \lfloor \mu_1(H) \rfloor + 1$. By the Cauchy interlacing theorem, we can further yield that

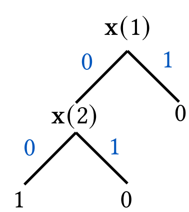
$$\chi(H) \leq \lfloor \mu_1(H) \rfloor + 1 \leq \lfloor \mu_1(G) \rfloor + 1.$$

Note that the number of neighbors of v is no larger than $\mu_1(G)$. Therefore, $\lfloor \mu_1(G) \rfloor + 1$ colors are enough to construct a proper coloring. □

2 *Sensitivity Conjecture*

2.1 *Boolean Function and Its Sensitivity*

Let $f\colon \{-1,1\}^n \to \{0,1\}$ be a boolean function. We can use a decision tree to describe f and the depth of the tree measures the complexity of f. For example, consider the function f that $f(\mathbf{x}) = 1$ iff $\mathbf{x}(1) = \mathbf{x}(2) = 0$. The decision tree for this function is



In contrast, when $f(\mathbf{x}) = \begin{cases} 1, & \text{if } \left(\sum_i \mathbf{x}(i)\right) \mod 2 = 1 \\ 0, & o.w. \end{cases}$, the depth of the decision tree is n, which means the function is much more complex.

There is another way to measure the complexity of f. We can find a polynomial p such that $p(\mathbf{x}) = f(\mathbf{x})$ for any $\mathbf{x} \in \{-1,1\}^n$. Larger degree of p indicates that f is more complex.

For every $\mathbf{x} \in \{-1,1\}^n$, let

$$s(f, \mathbf{x}) = |\{i \mid f(\mathbf{x}) \neq f(\mathbf{x} \oplus i)\}|$$

where $\mathbf{x} \oplus i$ means flipping the i-th bit of $\mathbf{x}$. The sensitivity of f is defined as $s(f) := \max_{\mathbf{x} \in \{-1,1\}^n} s(f, \mathbf{x})$. Just as indicated by its name, this quantity reflects the sensitivity to perturbations of f.

Similarly, we can define block sensitivity. Define $bs(f, \mathbf{x})$ as the maximum number t of disjoint subsets $B_1, \ldots, B_t \subseteq [n]$ such that $f(\mathbf{x}) \neq f(x \oplus B_i)$ for every $i \in [t]$. The block sensitivity of f is $bs(f) := \max_{\mathbf{x} \in \{-1,1\}^n} bs(f, \mathbf{x})$. Obviously, $s(f) \leq bs(f)$.

2.2 Sensitivity Conjecture

The sensitivity conjecture states that there exists positive constant c such that $bs(f) \leq (s(f))^c$. In other words, $bs(f)$ and $s(f)$ are equivalent in a polynomial sense.

Let Q_n be the n-dimensional hypercube[2]. In the work of [?] this conjecture has been reduced to the following proposition.

Proposition 4. *For any induced subgraph H of Q_n with $2^{n-1} + 1$ vertices, $\Delta(H) \geq \sqrt{n}$.*[3]

Hao Huang gave a remarkable proof of the above proposition and hence the sensitivity conjecture.

Proof. Let $A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $A_n = \begin{bmatrix} A_{n-1} & I \\ I & -A_{n-1} \end{bmatrix}$ for $n \geq 2$. We claim that $A_n^2 = nI$. This can be proved by induction. When $n = 1$, it is trivial to have $A_1^2 = I$. For $n \geq 2$,

$$A_n^2 = \begin{bmatrix} A_{n-1} & I \\ I & -A_{n-1} \end{bmatrix}^\top \begin{bmatrix} A_{n-1} & I \\ I & -A_{n-1} \end{bmatrix} = \begin{bmatrix} A_{n-1}^2 + I & 0 \\ 0 & A_{n-1}^2 + I \end{bmatrix} = nI.$$

This indicates that the eigenvalue of A_n is either $\sqrt{n}$ or $-\sqrt{n}$. Note that the trace of A_n is 0. Therefore,

$$\lambda_1 = \cdots = \lambda_{2^{n-1}} = \sqrt{n} \quad \text{and} \quad \lambda_{2^{n-1}+1} = \cdots = \lambda_{2^n} = -\sqrt{n}.$$

It can be verified that A_n is a signed adjacent matrix of Q_n. That is, $A_n(x, y) \neq 0$ indicates that $Q_n(x, y) \neq 0$ and $A_n(x, y) \in \{0, \pm 1\}$. Pick the columns and rows in H, we can get a principal submatrix of A_n, denoted as $A_n(H)$. Then we have $\Delta(H) = \|A_n(H)\|_\infty \geq \lambda_1(A_n(H))$. By the Cauchy interlacing theorem, $\lambda_1(A_n(H)) \geq \lambda_{2^{n-1}}(A_n) = \sqrt{n}$. This completes the proof. □

[2] A hypercube $Q_n = (V, E)$ is a graph with $V = \{0, 1\}^n$ and $E = \{\{x, y\} \mid x = y \oplus i \text{ for some } i\}$

[3] $\Delta(H)$ is the maximum degree of H.

The $2^{n-1}+1$ vertices can not be reduced any more in Proposition 4. Note that hypercube is bipartite. So there exists an independent set with 2^{n-1} vertices in Q_n.