Semidefinite programming and eigenvalues

Péter Hajnal

2024. Fall

Eigenvalues

Eigenvalue Problem

Given $M \in \mathcal{S}^n \subset \mathbb{R}^{n\times n}$ symmetric matrix. // Thus we know its eigenvalues are real.

Let's determine its eigenvalues:

$$\lambda_{\mathsf{max}} = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n = \lambda_{\mathsf{min}}.$$

The eigenvalues (with multiplicities) form the spectrum of the matrix.

J

Consider the all-1 matrix:

$$J = J_k = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \ddots & & 1 \\ \vdots & & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}_{k\times k}.$$

What are its eigenvalues, eigenvectors?

The Solution

• The all-1 vector: $\underline{1} = j = (1, 1, \cdots, 1)^{\mathsf{T}} \in \mathbb{R}^k$ is an eigenvector of the matrix: $J\underline{1} = k\underline{1}$. So, the eigenvalue associated with the found eigenvector is k.

• Consider another eigenvalue and its eigenvector, let it be: $(x_1, \cdots, x_k)^{\mathsf{T}} \in \mathbb{R}^k$.

• We know that it is orthogonal to $\underline{1}$, i.e., $\sum x_i = 0$.

• This idea is reversible: If $\sum x_i = 0$, then $J\underline{x} = \underline{0} = 0\underline{x}$.

• Thus, we found that 0 is an eigenvalue and certain associated eigenvectors span a $(k-1)$-dimensional subspace. So, 0 is at least $k-1$ times an eigenvector.

• With this, we have all eigenvalues:

$$k \geq 0 \geq 0 \geq \ldots \geq 0 \geq 0.$$

Specifically, $\lambda_{\max}(J_{k\times k}) = k$.

Maximal Eigenvalue Problem

Simplified Eigenvalue Problem

Given $M \in \mathcal{S}^n$, determine its maximal eigenvalue.

It is easy to see and well-known that

$$\lambda_{\max} = \max_{x \in \mathbb{R}^n : x \neq 0} \frac{x^\mathsf{T} M x}{x^\mathsf{T} x} = \max_{x \in \mathbb{R}^n : \|x\| = 1} x^\mathsf{T} M x.$$

So, determining the maximal eigenvalue can be formulated as:

Maximize	$x^\mathsf{T} Mx$-t
subject to	$\|x\| = 1.$

Maximal Eigenvalue Otherwise

Observation

The eigenvalues of $\lambda I - M$ are

$$\lambda - \lambda_n \geq \lambda - \lambda_{n-1} \geq \ldots \geq \lambda - \lambda_1.$$

• These eigenvalues are nonnegative if and only if $\lambda \geq \lambda_1 = \lambda_{\max}$. The nonnegativity of eigenvalues of $\lambda I - M$ precisely means the positive semidefiniteness of $\lambda I - M$.

• The smallest such λ value is $\lambda_{\max}$. Our reformulation:

Minimize	λ-t
subject to	$\lambda I - M \succeq 0$

• This is an SDP formulation of determining $\lambda_{\max}$.

A More Complex Eigenvalue Problem

Let $X = \sum_{1=1}^{n} x_i A_i$, where $A_i \in \mathcal{S}^k$ (so $X \in \mathcal{S}^k$ holds).

That is, X has the following form

$$\begin{pmatrix} \alpha_{11}^{(1)} x_1 + \alpha_{11}^{(2)} x_2 + \dots & \alpha_{12}^{(1)} x_1 + \alpha_{12}^{(2)} x_2 + \dots & \cdots & \alpha_{1n}^{(1)} x_1 + \alpha_{1n}^{(2)} x_2 + \dots \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1}^{(1)} x_1 + \alpha_{11}^{(2)} x_2 + \dots & \alpha_{n2}^{(1)} x_1 + \alpha_{n2}^{(2)} x_2 + \dots & \cdots & \alpha_{nn}^{(1)} x_1 + \alpha_{nn}^{(2)} x_2 + \dots \end{pmatrix}_{n\times}$$

i.e., an $n \times n$ matrix, where each element is a linear function.

Task

Find $x \in \mathbb{R}^n$ such that $\lambda_{\max}(X)$ is minimized.

SDP Formulation

The „more complex" task can be easily rephrased with previous ideas:

Minimize	μ-t
subject to	$X = \sum_{i=1}^{n} x_i A_i$
	$\mu I - X \succeq 0.$

That is, our eigenvalue question can again be formulated as an SDP problem.

The PLAN

- In the following, we tackle a difficult ($\mathcal{NP}$-complete) graph-theoretical problem.

- We provide an estimation for its optimum using eigenvalues.

- Then, we formulate the assertion of the best estimate as an SDP problem.

Break

Reminder: Adjacency Matrix

The adjacency matrix A_G of a simple graph G

$$A_G = \begin{array}{c} \\ \\ y \\ \\ \vdots \\ x \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccc} \quad & y & \quad & x & \quad \end{array} \\ \begin{pmatrix} 0 & & \cdots & & \\ & 0 & \cdots & \begin{cases} 1, & \text{if } yx \in E \\ 0, & \text{otherwise} \end{cases} & \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ & \begin{cases} 1, & \text{if } xy \in E \\ 0, & \text{otherwise} \end{cases} & \cdots & 0 & \\ & & \cdots & & 0 \end{pmatrix} \end{array}.$$

- A_G can be imagined as being indexed by the vertices V, with $n = |V|$.
- To represent A_G, we need to fix an ordering of the vertices.
- In an ordering-independent view of the matrix, we must interpret it as a function $V \times V \to \{0, 1\}$.

Reminder: Independent Sets

Definition

$F \subset V(G)$ is an independent set if every edge has at most one endpoint in F. That is, there are no edges within F. That is, $G|_F$ is an empty graph.

For a given simple graph G, the associated optimization problem is

Maximize	$\|F\|$-t
subject to	F independent set

The standard notation for the optimal value is $\alpha(G)$.

Independent Sets and Adjacency Matrix

If F is an independent set, then the vertices in F indicate a zero submatrix in A_G:

$$\begin{array}{cc} & \overbrace{\hphantom{0 \;\; \ldots \;\; 0}}^{F} \\ F\left\{\vphantom{\begin{matrix}0\\ \vdots\\ 0\end{matrix}}\right. & \begin{pmatrix} & & & & \\ & & & & \\ & 0 & \ldots & 0 & \\ & \vdots & \ddots & \vdots & \\ & 0 & \ldots & 0 & \\ & & & & \end{pmatrix} \end{array}$$

i.e., if we remove the rows/columns corresponding to vertices outside F, we get a matrix filled with zeros. That is, $A_G|_{F \times F} \equiv 0$.

A Twist

For technical reasons, let's consider another formulation of the problem, where we work with $\overline{A_G}$ instead of A_G.

In this matrix, we interchange the 0s and 1s: the main diagonal contains 1s, elsewhere, we map non-edges between vertex pairs to 1, and edges to 0.

That is,

$$\overline{A_G} = \begin{array}{cc} & y \\ \begin{array}{c} \\ y \\ \\ x \\ \\ \end{array} & \begin{pmatrix} 1 & & \cdots & \\ & 1 & \cdots & \\ & & \ddots & \\ & \begin{cases} 0, & \text{if } xy \in E \\ 1, & \text{otherwise} \end{cases} & & \\ & & \cdots & 1 \end{pmatrix} \end{array} = J - A_G = I + A_{\overline{G}}.$$

F is an independent set if and only if $\overline{A_G}|_{F\times F} = J$, the constant matrix of ones.

The Idea

- Based on these, let's take the eigenvector corresponding to the $|F|$ eigenvalue of the $|F| \times |F|$ submatrix of $\overline{A_G}$ ($\in \mathbb{R}^{F \times F}$).
- Extend it with zeros to obtain a vector in $\mathbb{R}^V$, let's call it χ_F.
- It's easy to verify that

$$|F| = \frac{\chi_F^\mathsf{T} \overline{A_G} \chi_F}{\chi_F^\mathsf{T} \chi_F} \leq \max_{x : x \in \mathbb{R}^V - \{0\}} \frac{x^\mathsf{T} \overline{A_G} x}{x^\mathsf{T} x} = \lambda_{\max}(\overline{A_G}).$$

- We arrive at the following result:

Theorem

Let G be a simple graph, F an independent set in our graph. Then

$$\lambda_{\max}\left(\overline{A_G}\right) \geqslant |F|.$$

Specifically,

$$\lambda_{\max}\left(\overline{A_G}\right) \geqslant \alpha(G).$$

Why is this useful?

• From complexity theory, we know that determining the size of the largest independent set is an $\mathcal{NP}$-hard problem.

• From numerical methods, we know that the maximum eigenvalue can be efficiently determined.

• By calculating $\lambda_{\max}\left(\overline{A_G}\right)$, we obtain an estimate for an $\mathcal{NP}$-hard function.

Squeezing the Idea

Observation

In this line of thought, we only used the property that in $\overline{A_G}$, both the main diagonal and the non-edges have 1s.

Consequence

Let G be any simple graph.

- Let $M \in \mathcal{S}^V$ be any arbitrary matrix that satisfies
  - (i) 1's on the main diagonal,
  - (ii) 1's for non-edges

// Let this property be T_G for $M \in \mathcal{S}^V$

- Let F be any independent set.

Then

$$\lambda_{\max}(M) \geqslant |F|.$$

Maximally Utilizing the Consequence

- The consequence has two „participants".

(i) the matrix with property T_G,

(ii) the independent set F.

- One is on the left, the other on the right side.
- Thus, it's straightforward to formulate the sharpest version of the theorem.

Theorem

Let G be a simple graph. Then

$$\min\{\lambda_{\max}(M) : \; M \in \mathcal{S}^V \text{ satisfies } T_G\} \geqslant \max\{|F| : \; F \text{ is an independent set}\}.$$

The Left-hand Side of the Inequality

Minimize	$\lambda_{\max}(M)$-t
subject to	$M_{uu} = 1$, for all $u \in V$,
	$M_{uv} = 1$, for all $uv \notin E$,
	$M \in \mathcal{S}^n$.

Let's reformulate the problem and see that its determination is an SDP problem.

Reformulation: Notations

Let $e = xy \in E(G)$ be any edge. Let S_e be the matrix where only the positions xy and yx contain 1, while everywhere else contains 0. That is,

$$S_e = \begin{array}{c c} & \begin{array}{cccccccc} & & & x & & y & & \end{array} \\ \begin{array}{c} \\ \\ \\ x \\ \\ y \\ \\ \\ \end{array} & \begin{pmatrix} 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \end{pmatrix} \end{array}.$$

That is,

$$S_e(u, v) = \begin{cases} 1, & u = x, v = y \text{ or } u = y, v = x \\ 0 & \text{otherwise.} \end{cases}$$

The Reformulated Problem

After this, a symmetric matrix M's satisfying the property T_G is equivalent to $M = J - \sum_{e \in E} x_e S_e$ for some $x_e \in \mathbb{R}^E$ vector: The main diagonal and the non-edges of the J matrix are 1, and everywhere else (at the position of edge e) we modify it by x_e (to any value).

Then, this is an SDP problem as in the first example:

Minimize	μ-t
subject to	$M = J - \sum_{e \in E} x_e S_e$
	$\mu I - M \succeq 0.$

That is (in the second normal form),

Minimize	μ-t
subject to	$-\mu I - \sum_{e \in E} x_e S_e \preceq -J.$

Summary

Summarizing:

Definition/Notation

Let G be a simple graph. Then

$$\vartheta(G) = \text{ the optimal value of the above SDP problem.}$$

We obtained that

Theorem

Let G be a simple graph. Then

$$\vartheta(G) \geq \alpha(G).$$

Knowing that the optimum of an SDP optimization problem can be efficiently determined, the left-hand side of the theorem's inequality is $\mathcal{NP}$-hard, while the right-hand side is tractable.

Break

Clique Cover Problem

Problem

Given a simple graph G. Cover its vertex set with as few cliques as possible.

Let $\overline{\chi}(G)$ be the minimum number of cliques required for this cover.

• A clique cover is a proper coloring of the complement graph. That is, the clique cover problem is the coloring problem for the complement graph, i.e., $\overline{\chi}(G) = \chi(\overline{G})$.
• Specifically, the clique cover problem is $\mathcal{NP}$-hard.
• For a simple graph G, we can again describe it with a matrix. For us, the $A_{\overline{G}}$ matrix will be „convenient".
• Take a clique cover of G. Let ℓ be the number of cliques. That is, $V = K_1 \dot\cup K_2 \dot\cup \ldots \dot\cup K_\ell$, where K_i are disjoint cliques.

Vectors, Matrices

We can reflect the classification of vertices in the linear algebraic notation.

• We think of the elements of $\mathbb{R}^V$ being divided into ℓ blocks:

$$(\overbrace{x_1,\dots}^{K_1}|\overbrace{\quad\dots\quad}^{K_2}|\dots|\overbrace{\dots,x_n}^{K_\ell}).$$

• Similarly, a matrix of type $V \times V$ can be considered as a block matrix of type $\ell \times \ell$:

$$M = \begin{array}{c c} & \begin{array}{cccc} \overbrace{\quad\quad}^{K_1} & \overbrace{\quad\quad}^{K_2} & \dots & \overbrace{\quad\quad}^{K_\ell} \end{array} \\ \begin{array}{c} K_1\{ \\ K_2\{ \\ \vdots \\ K_\ell\{ \end{array} & \left(\begin{array}{c|c|c|c} \quad & \quad & \quad & \quad \\ \hline \quad & \quad & \quad & \quad \\ \hline \quad & \quad & \quad & \quad \\ \hline \quad & \quad & \quad & \quad \end{array}\right) \end{array}.$$

The Neighborhood Block Matrix

Let's specifically look at $A_{\overline{G}}$:

$$A_{\overline{G}} = \begin{array}{cc} & \begin{array}{cccc} \overbrace{\qquad}^{K_1} & \overbrace{\qquad}^{K_2} & \cdots & \overbrace{\qquad}^{K_\ell} \end{array} \\ \begin{array}{c} K_1\{ \\ K_2\{ \\ \vdots \\ K_\ell\{ \end{array} & \left(\begin{array}{c|c|c|c} \boxed{0} & & & \\ \hline & \boxed{0} & & \\ \hline & & \ddots & \\ \hline & & & \boxed{0} \end{array}\right) \end{array}.$$

Let the eigenvalues of $A_{\overline{G}}$ be denoted by
$\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n = \lambda_{\max}$ $(n = |V|)$.

v and $\widetilde{v}$

Let $v^{\mathsf{T}} = (v_1|v_2|\dots|v_l)$ be the eigenvector corresponding to $\lambda_{\max}$.

Define

$$\widetilde{v}^{\mathsf{T}} := \Big(\|v_1\|, 0, \dots, 0\Big|\|v_2\|, 0, \dots, 0\Big|\dots\Big|\|v_\ell\|, 0, 0, \dots, 0\Big)$$

where $\|w\| = \sqrt{\sum_{i=1}^{d} (w_i)^2}$, the L^2 norm of the d-dimensional vector w.

Q Orthogonal Matrix

- Clearly, $\|\tilde{v}\| = \|v\|$, so there exists an orthogonal matrix Q that transforms $\widetilde{v}$ to v: $Qv = \widetilde{v}$.
- Q can be chosen to be *compatible* with the existing blocking: the L^2 norms of vectors $\widetilde{v}$ and v within the blocks are preserved.
- Thus, there exist orthogonal matrices Q_i such that $Q_i\widetilde{v}_i = v_i$, where $\widetilde{v}_i = (\|v_i\|_2, 0, 0, \ldots, 0)^\mathsf{T} \in \mathbb{R}^{K_i}$.
- Then

$$Q = \begin{array}{r} \\ K_1\{ \\ K_2\{ \\ \vdots \\ K_\ell\{ \end{array} \begin{array}{c} \begin{array}{cccc} \overbrace{}^{K_1} & \overbrace{}^{K_2} & \ldots & \overbrace{}^{K_\ell} \end{array} \\ \left(\begin{array}{c|c|c|c} Q_1 & \boxed{0} & \ldots & \boxed{0} \\ \hline \boxed{0} & Q_2 & \ldots & \boxed{0} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \boxed{0} & \boxed{0} & \ldots & Q_\ell \end{array}\right) \end{array}, \qquad \text{and} \quad Q\widetilde{v} = v.$$

Observation

Observation

If u is an eigenvector corresponding to the eigenvalue λ of $A_{\overline{G}}$, then $Q^{-1}u = Q^{\mathsf{T}}u$ is an eigenvector of the matrix $Q^{-1}A_{\overline{G}}Q$ corresponding to the same eigenvalue λ.

Indeed,

$$(Q^{-1}A_{\overline{G}}Q)(Q^{-1}u) = Q^{-1}A_{\overline{G}}(QQ^{-1}u) = Q^{-1}A_{\overline{G}}u = Q^{-1}\lambda u = \lambda Q^{-1}u.$$

$Q^{-1}A_{\overline{G}}Q$

Specifically, the eigenvalues of $A_{\overline{G}}$ and $Q^{-1}A_{\overline{G}}Q$ coincide.

Note that the block structure and the zero blocks on the main diagonal of $Q^{-1}A_{\overline{G}}Q$ can also be recognized:

$$Q^{-1}A_{\overline{G}}Q = \begin{array}{c c} & \begin{array}{cccc} \overbrace{\quad}^{K_1} & \overbrace{\quad}^{K_2} & \cdots & \overbrace{\quad}^{K_\ell} \end{array} \\ \begin{array}{c} K_1\{ \\ K_2\{ \\ \vdots \\ K_\ell\{ \end{array} & \left(\begin{array}{c|c|c|c} \boxed{0} & & & \\ \hline & \boxed{0} & & \\ \hline & & \ddots & \\ \hline & & & \boxed{0} \end{array}\right) \end{array}.$$

$R = A_G|_{F \times F}$

• That is, $\lambda_{\max}$ is an eigenvalue of the matrix $Q^{-1}A_{\overline{G}}Q$, and $Q^{-1}v = \widetilde{v}$ is the corresponding eigenvector to $\lambda_{\max}$.

• This eigenvector contains at most ℓ non-zero coordinates (the first element of each block may be non-zero).

• Let F be the set of vertices corresponding to the first elements of the blocks. That is, we take one vertex from each color class (considering the complement graph).

• The set F can be seen as a set of rows or columns.

• We form a submatrix R of $A_{\overline{G}}$ by removing the rows and columns not in F. (This operation is called symmetric submatrix extraction.)

Cauchy's Theorem

Let $M_{s\times s}$ be a symmetric matrix, and let's form a symmetric submatrix $R_{t\times t}$. Let $\lambda_1 \leq \ldots \leq \lambda_s$ be the eigenvalues of $M_{s\times s}$, and let $\mu_1 \leq \ldots \leq \mu_t$ be the eigenvalues of $R_{t\times t}$. Then

(i) $\lambda_1 \leq \mu_1$, $\lambda_2 \leq \mu_2$, ..., $\lambda_t \leq \mu_t$,

(ii) $\mu_t \leq \lambda_s$, $\mu_{t-1} \leq \lambda_{s-1}$, ..., $\mu_1 \leq \lambda_{s-t+1}$.

Where Are We?

- We obtained a $\ell \times \ell$ matrix R of type $F \times F$ from $Q^{-1}A_{\overline{G}}Q$ through symmetric submatrix extraction.
- It's easy to see that $\widetilde{v}|_F \in \mathbb{R}^F$ (vector formed by the first elements of the blocks of $\widetilde{v}$) is an eigenvector of R.
- Moreover, the corresponding eigenvalues are λ_{max}. Let's specialize Cauchy's theorem to our case.
- The eigenvalues of $Q^{-1}A_{\overline{G}}Q$ (and $A_{\overline{G}}$): $\lambda_{\text{min}} = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n = \lambda_{\text{max}}$.
- What does Cauchy's theorem say about the eigenvalues of R ($\mu_1 \leq \mu_2 \leq \ldots \leq \mu_\ell$)? They lie between λ_{min} and λ_{max}.
- Based on the above, the largest eigenvalue is λ_{max}.
- The sum of eigenvalues is the trace of our matrix, which is 0.

Where Are We? Summarized

- If we assume that the known eigenvalues of $\lambda_{\max}$ are accompanied by $\ell - 1$ other eigenvalues estimated to be $\lambda_{\min}$, then we get: $0 = \text{trace } R = \sum_{i=1}^{\ell} \mu_i \geq (\ell - 1)\lambda_{\min} + \lambda_{\max}$.
- Hence

$$-(\ell - 1)\lambda_{\min} \geq \lambda_{\max},$$

If $A_{\overline{G}}$ has eigenvalues other than 0, then $\lambda_{\min} \lneqq 0 \lneqq \lambda_{\max}$ (we know their sum is 0). This occurs when $\overline{G}$ is not an empty graph, meaning G is not complete.

- Then we can divide by $-\lambda_{\min}$:

$$\ell - 1 \geq \frac{\lambda_{\max}}{-\lambda_{\min}}.$$

- That is,

$$\ell \geq 1 + \frac{\lambda_{\max}}{-\lambda_{\min}}.$$

Hoffman's Theorem, Complementary Form

Hoffman's Theorem, Complementary Form

Let G be a non-complete graph.

$A_{\overline{G}}$ is the adjacency matrix of the complement graph, i.e., it has 0's on its main diagonal, and off-diagonal 1's encode non-adjacency of G (or adjacency of 0's in G).

Let ψ be a clique cover, $\ell(\psi)$ be the number of cliques.

Then

$$\ell(\psi) \geq 1 + \frac{\lambda_{\max}(A_{\overline{G}})}{-\lambda_{\min}(A_{\overline{G}})}.$$

Hoffman's Theorem

If we apply the theorem to the complement of G, we obtain an estimation for the chromatic number.

Hoffman's Theorem, Original/Coloring Form

Let G be a simple graph, not empty. Let A_G be the adjacency matrix of the graph.

$$\chi(G) \geq 1 + \frac{\lambda_{\text{max}}(A_G)}{-\lambda_{\text{min}}(A_G)}.$$

Extracting the „Essence"

The proof only depended on the fact that the main diagonal of our matrix $A_{\overline{G}}$ consists of 0's and the off-diagonals encode 0's for adjacency.

Definition

Let $\widetilde{T}_G$ be a $V \times V$ symmetric matrix having 0's on its main diagonal and encoding the 0's for adjacency in G.

- The above proof can be repeated with $\widetilde{T}_G$ instead of $A_{\overline{G}}$.
- This opens up the possibility to improve the Hoffman's estimation, even optimizing these improvements.

Hoffman's Theorem, Strong Form

Hoffman's Theorem, Complementary Strong Form

Let G be a simple graph that is not a complete graph.

Let M be a symmetric matrix of type $V \times V$ with property $\widetilde{T}_G$.

Let f be a clique cover, and $\ell(f)$ be the number of cliques in the clique cover.

Then

$$\ell(f) \geq 1 + \frac{\lambda_{\mathrm{max}}(M)}{-\lambda_{\mathrm{min}}(M)}.$$

That is,

$$\min\{\ell(f) : \ f \text{ is a clique cover}\} \geq \max\left\{1 + \frac{\lambda_{\mathrm{max}}(M)}{-\lambda_{\mathrm{min}}(M)} : \ M \text{ has property } \widetilde{T}_G\right\}.$$

The value of the left-hand side of the final inequality is $\overline{\chi}(G)$.

Break

The Right-hand Side

Let (L) denote the optimization problem on the right-hand side:

$$1 + \frac{\lambda_{\max}(M)}{-\lambda_{\min}(M)}$$

subject to:

$$M_{uu} = 0 \text{ for all } u \in V$$

and

$$\langle M, S_e \rangle = 0 \text{ for all } e \in E$$

with $M \in \mathcal{S}^n$.

This problem can be reformulated into an SDP form.

The journey towards our "beloved" SDP form will be long.

Intermediate Problem

Let (K) be an intermediate problem:

$$\lambda_{\max}(N)$$

subject to:

$$N_{uu} = 1 \text{ for all } u \in V$$

and

$$N_{uv} = 0 \text{ for all } uv \in E$$

with $N \succeq 0$.

Theorem

The optimization problems describing the derivation of Hoffman's Theorem, (L), and the intermediate problem (K) are equivalent.

The claim is that the optimal values of the two problems coincide. This is demonstrated by establishing both directions of the inequality between them.

From (L) Solution to (K) Solution with Non-worsening Objective

Maximize	$1 + \frac{\lambda_{\max}(M)}{-\lambda_{\min}(M)}$-t
subject to	$M_{uu} = 0$ for all $u \in V$
	$\langle M, S_e \rangle = 0$ for all $e \in E$
	$M \in \mathcal{S}^n$.

We construct matrix $N = I + \frac{1}{-\lambda_{\min}(M)} M$ from a feasible M.

It can be seen that the constructed N is a feasible solution to the intermediate problem.

The smallest eigenvalue of $\frac{1}{-\lambda_{\min}(M)} M$ will be -1. Thus, adding the identity matrix ensures all eigenvalues become non-negative.

Moreover, its objective function value remains the same as the one defined on M in the original problem.

From (K) Solution to (L) Solution with Non-worsening Objective

Maximize	$\lambda_{\max}(N)$-t
subject to	$N_{uu} = 1$ for all $u \in V$
	$N_{uv} = 0$ for all $uv \in E$
	$N \succeq 0.$

This train of thought can be reversed. Let N be an optimal solution to the intermediate problem (K).

Observation

Suppose an off-diagonal element of a positive semidefinite matrix A is 0. Then, its corresponding row and column vectors are all 0 vectors.

From (K) Solution to (L) Solution (continued)

• If N has a 0 off-diagonal element, then it has 0 row and column vectors on that part. On this part, let N be 1 on the diagonal and 0 otherwise. The essential part of the construction lies in defining the other elements. We focus on this: essentially, we assume that the off-diagonal elements of Λ are non-zero.

• Let u be a vector formed from square roots of the positive (assumed) elements on the diagonal of Λ. This will obviously be a unit vector.

• Let $U = uu^{\mathsf{T}}$. According to our assumption, this is a non-zero matrix.

• Let N be the matrix for which $N \cdot_H U = \Lambda$ holds.

• This is a feasible solution to problem (K) with a non-worsening objective function.

The Final Form, $(\widetilde{L})$

After these, we show that the intermediate problem can be reformulated into the following SDP form $(\widetilde{L})$:

Maximize	$\langle J, \Lambda\rangle$-t
subject to	$\langle S_e, \Lambda\rangle = 0$
	$\langle I, \Lambda\rangle = 1$
	$\Lambda \succeq 0.$

Theorem

The intermediate problem (K) and the SDP problem $(\widetilde{L})$ are equivalent (their optimal values coincide).

Corollary

The optimal value of the SDP problem $(\widetilde{L})$ is an estimate of the twisted Hoffman bound for the clique cover problem.

From (K) Solution to $(\widetilde{L})$ Solution with Non-worsening Objective

The value of the objective function is

$$\langle J, \Lambda \rangle = \langle J, N \cdot_H (uu^\mathsf{T}) \rangle = u^\mathsf{T} N u = \lambda_{\max}(N).$$

If we examine the positions in $\Lambda = N \cdot_H U$ where 0 appears in N, by definition, we see 0s.

$$\langle I, \Lambda \rangle = Tr(N \cdot_H (uu^\mathsf{T})) = Tr(uu^\mathsf{T}) = u^\mathsf{T} u = 1.$$

Λ is Positive Semidefinite

• The positive semidefiniteness of Λ is left to be shown. This is obvious from the following lemma.

Lemma

Let $A, B \succeq 0$. Then, $A \cdot_H B \succeq 0$.

• The lemma easily follows from the fact that any positive semidefinite matrix can be written as a sum of positive semidefinite matrices of rank 1.

• By writing A and B in this form, we see that the parentheses in $A \cdot_H B$ can be expanded, resulting in $A \cdot_H B$ being the Hadamard product of rank 1 positive semidefinite matrices.

• However, the positive semidefiniteness of these is obvious.

• With this lemma and one direction of the inequality, we have derived both directions.

From $(\widetilde{L})$ Solution to (K) Solution

Let's reverse the above argument for the other direction of the inequality.

Maximize	$\langle J, \Lambda\rangle$-t
subject to	$\langle S_e, \Lambda\rangle = 0$
	$\langle I, \Lambda\rangle = 1$
	$\Lambda \succeq 0.$

- Suppose $\Lambda = \frac{1}{|V|} \cdot I$ is a feasible solution to the SDP. The objective function value is at least 1.
- Let Λ be any feasible solution to the SDP.

Note

Assume that one of the diagonal elements of a positive semidefinite matrix is 0. Then, the row and column vectors on this element are all 0 vectors.

From $(\widetilde{L})$ Solution to (K) Solution (continued)

• If Λ has a 0 diagonal element, then it has 0 row and column vectors on that part. On this part, let N be such that it is 1 on the diagonal and 0 otherwise. The essential part of the construction lies in defining the other elements. We focus on this: essentially, we assume that the off-diagonal elements of Λ are non-zero.

• Let u be a vector formed from square roots of the positive (assumed) elements on the diagonal of Λ. This will obviously be a unit vector.

• Let $U = uu^\mathsf{T}$. According to our assumption, this is a non-zero matrix.

• Let N be the matrix for which $N \cdot_H U = \Lambda$ holds.

• This is a feasible solution to problem (K) with a non-worsening objective function.

Summary

Definition

Let $\widetilde{\vartheta}(G)$ be the optimum of the $(\widetilde{L})/(\mathrm{K}),(\mathrm{L})$ optimization problem.

Our reformulations stemmed from formalizing the clique cover problem as a twisted Hoffman bound estimation. Thus, we obtain the following theorem.

Theorem

$$\widetilde{\vartheta}(G) \leq \overline{\chi}(G).$$

Reminder

For both $\alpha(G)$ and $\overline{\chi}(G)$, we provided an SDP problem, the optimal value of which estimates the corresponding graph parameter:

Minimize	μ-t
subject to	$-\mu I - \sum_{e\in E} x_e S_e \preceq -J.$

Maximize	$\langle J, \Lambda\rangle$-t
subject to	$\langle S_e, \Lambda\rangle = 0$
	$\langle I, \Lambda\rangle = 1$
	$\Lambda \succeq 0.$

Summary: The Sandwich

Observation

(i) The two SDP problems are dual to each other.

(ii) Both SDP problems satisfy conditions that ensure strong duality.

Thus, the two concepts coincide.

Definition: The $\vartheta(G)$ of Graph G by Lovász

$$\vartheta(G) = \widetilde{\vartheta}(G).$$

The two previous estimates are summarized by the following theorem.

Lovász's Sandwich Theorem

$$\alpha(G) \leq \vartheta(G) \leq \overline{\chi}(G).$$

The Moral

• The intermediate function initially appears very complex, unnatural.

• The two extreme functions have elementary definitions, understandable even to an interested high school student.

• However, the two extreme graph optimization questions are complex. $\mathcal{NP}$-hard. We see no possibility for efficient calculation (according to the general belief).

• The value of the intermediate function, however, can be calculated/approximated efficiently (SDP problems are manageable).

This is the End!

Thank you for your attention!