Aggregate Bounds on the Eigenvalues of Principal Submatrices and Majorization Relations

Mengxu Yuan
(joint work with Professor Hristo Sendov)

Department of Mathematics
Western University
London, Ontario, Canada

February 6, 2026

LE MENU MATHÉMATIQUE

An intellectual gastronomic experience

Bon appétit mathématique!

Apéritif

Majorization Theory

The concept of Majorization is central in matrix analysis

For $x, y \in \mathbb{R}^N$, let $x^\downarrow, y^\downarrow$ be sorted in non-increasing order

Definition (Majorization, $\succ$)

We say x majorizes y, denoted by $x \succ y$, if

$$\sum_{i=1}^{k} x_i^\downarrow \geq \sum_{i=1}^{k} y_i^\downarrow \quad (k < N) \quad \text{and} \quad \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} y_i$$

How to visualize $x \succ y$?

Think of y as a “smoothed” or “averaged” version of x

Examples of majorisation

Example

Let $x = (5, 3, 1)$, $y = (4, 3, 2)$
Partial sums of x: $\mathbf{5}, 5 + 3 = \mathbf{8}, 5 + 3 + 1 = \mathbf{9}$
Partial sums of y: $\mathbf{4}, 4 + 3 = \mathbf{7}, 4 + 3 + 2 = \mathbf{9}$
$5 \geq 4$, $8 \geq 7$, and $9 = 9$. Thus $x \succ y$

Example

For $x \in \mathbb{R}^n_+$ such that $\sum x_i = 1$, then

$$(1, 0, \dots, 0) \succ x \succ \left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right)$$

Majorization is a preorder

Majorization is a preorder on $\mathbb{R}^n$, because it satisfies
Reflexivity : $x \succ x$
Transitivity : $x \succ y$ and $y \succ z \implies x \succ z$

Missing piece : antisymmetry
Partial order requires : $x \succ y$ and $y \succ x \implies x = y$
Majorization fails this
If we restrict the domain to sorted vectors, then $\succ$ becomes a partial order

Example

Let $x = (1, 0)$, $y = (0, 1)$, then

$$x^{\downarrow} = (1, 0) = y^{\downarrow} \implies x \succ y \text{ and } y \succ x$$

But clearly $x \neq y$. They are distinct, yet equivalent under $\succ$

Doubly Stochastic Matrices

Definition (Doubly stochastic matrix, S)

$S = (s_{ij}) \in \mathbb{R}^{n \times n}$ is doubly stochastic if $s_{ij} \geq 0, \sum_{j=1}^{n} s_{ij} = 1, \sum_{i=1}^{n} s_{ij} = 1$

Theorem (Birkhoff-von Neumann theorem)

The set of doubly stochastic matrices is the convex hull of permutation matrices

$S = \sum_{k=1}^{N} \theta_k P_k$, *where* $\theta_k \geq 0, \sum \theta_k = 1$

Theorem (Hardy-Littlewood-Pólya)

For $x, y \in \mathbb{R}^n$, *the following are equivalent*

1. $x \succ y$
2. $y = Sx$ *for some doubly stochastic matrix* S
3. $y \in conv\{Px : P \text{ is a permutation matrix}\}$

Schur's Theorem

Eigenvalues of A is denoted by $\lambda(A) = (\lambda_1, \ldots, \lambda_n)$ with $\lambda_1 \geq \cdots \geq \lambda_n$

Schur's Theorem

For any $n \times n$ Hermitian matrix A

$$\lambda(A) \succ \mathrm{diag}(A)$$

Since diagonal entries are just 1×1 principal submatrices
Schur's theorem relates spectrum of A to its principal submatrices of size 1

Hadamard's inequality

A function $f : \mathbb{R}^n \to \mathbb{R}$ is Schur-convex if $x \succ y \implies f(x) \geq f(y)$
f is Schur-concave if $x \succ y \implies f(x) \leq f(y)$

Proposition (Schur-concavity of product)

If $x, y \in \mathbb{R}^n_+$ (non-negative vectors) and $x \succ y$, then

$$\prod_{i=1}^{n} x_i \leq \prod_{i=1}^{n} y_i$$

Example (Hadamard's inequality)

Let A be positive semidefinite. By Schur's theorem, $\lambda(A) \succ \mathrm{diag}(A)$. Then

$$\det(A) = \prod_{i=1}^{n} \lambda_i(A) \leq \prod_{i=1}^{n} A_{ii}$$

Cauchy's interlacing theorem

Aim to generalize relationships between A and its submatrices of any size
Let A be an $n \times n$ Hermitian matrix
Let A_k be the principal submatrix of A by deleting the k-th row and column
Eigenvalues of A_k is denoted by $\lambda(A_k) = (\mu_{k,1}, \ldots, \mu_{k,n-1})$

Cauchy's interlacing theorem

The eigenvalues of A and *any individual* principal submatrix A_k satisfy

$$\lambda_1 \geq \mu_{k,1} \geq \lambda_2 \geq \mu_{k,2} \geq \cdots \geq \mu_{k,n-1} \geq \lambda_n$$

Thompson's bounds

Cauchy's interlacing theorem describes one submatrix at a time
What about the aggregate behavior of all submatrices $A_1, \ldots, A_n$?

Theorem (Thompson, 1966)

For any fixed j ($1 \leq j \leq n-1$), sum of j-th eigenvalues satisfies

$$\lambda_j + (n-1)\lambda_{j+1} \leq \sum_{k=1}^{n} \mu_{k,j} \leq (n-1)\lambda_j + \lambda_{j+1}$$

Can we extend Thompson's bounds to sums over a range of indices $[\ell, r]$?

$$\sum_{k=1}^{n} \sum_{j=\ell}^{r} \mu_{k,j} \quad \text{vs.} \quad \text{Sums of } \lambda$$

Our first main result

We derived aggregate bounds for the sum of eigenvalues over $[\ell, r]$

Theorem (H.S., M.Y.)

For any $1 \leq \ell \leq r \leq n-1$

$$(r-\ell+1)\lambda_\ell + (n-1)\sum_{j=\ell}^{r}\lambda_{j+1} \leq \sum_{k=1}^{n}\sum_{j=\ell}^{r}\mu_{k,j} \leq (n-1)\sum_{j=\ell}^{r}\lambda_j + (r-\ell+1)\lambda_{r+1}$$

Our bounds are tighter

Let's try naive approach
Simply summing Thompson's upper bound for $j = \ell, \dots, r$

$$\text{Naive UB} = \sum_{j=\ell}^{r} [(n-1)\lambda_j + \lambda_{j+1}] = (n-1)\sum_{j=\ell}^{r} \lambda_j + \sum_{j=\ell}^{r} \lambda_{j+1}$$

Our new aggregate bound

$$\text{Our UB} = (n-1)\sum_{j=\ell}^{r} \lambda_j + (r-\ell+1)\lambda_{r+1}$$

Let's subtract our bound from the naive bound

$$\text{Naive} - \text{Ours} = \sum_{j=\ell}^{r} \lambda_{j+1} - (r-\ell+1)\lambda_{r+1} = \sum_{j=\ell}^{r} (\lambda_{j+1} - \lambda_{r+1}) \geq 0$$

A corollary and reprove Johnson, Robinson's result

This theorem implies the following corollary

Corollary

$$(n-\ell)\lambda_\ell + (n-1)\sum_{j=\ell+1}^{r}\lambda_j + r\lambda_n \leq \sum_{k=1}^{n}\sum_{j=\ell}^{r}\mu_{k,j} \leq (n-\ell)\lambda_1 + (n-1)\sum_{j=\ell+1}^{r}\lambda_j + r\lambda_{r+1}$$

This corollary reproves a known result

Theorem (Johnson, Robinson, 1981)

$$\max_{1\leq k\leq n}\mu_{k,j} \geq \frac{n-j}{n}\lambda_j + \frac{j}{n}\lambda_n \text{ and } \min_{1\leq k\leq n}\mu_{k,j} \leq \frac{n-j}{n}\lambda_1 + \frac{j}{n}\lambda_{j+1}$$

Our second main result

In order to get aggregate bounds, we need the following theorem
Let $\lambda_1 \geq \cdots \geq \lambda_n$ be reals and $\omega_i \geq 0$ be weights $(\sum_{i=1}^n \omega_i = 1)$
Let $\mu_1 \geq \cdots \geq \mu_{n-1}$ be roots of

$$p(x) = \sum_{i=1}^n \omega_i \prod_{\substack{j=1 \\ j \neq i}}^n (x - \lambda_j)$$

For any $[\ell, r]$, we derived the following bounds

Theorem (H.S., M.Y.)

$$\sum_{j=\ell}^r \lambda_{j+1} + \sum_{j=\ell}^r \frac{\omega_{j+1}}{L_{r+1}}(\lambda_\ell - \lambda_{j+1}) \leq \sum_{j=\ell}^r \mu_j \leq \sum_{j=\ell}^r \lambda_j - \sum_{j=\ell}^r \frac{\omega_j}{U_\ell}(\lambda_j - \lambda_{r+1})$$

where $U_\ell := \sum_{i=\ell}^n \omega_i$ *and* $L_{r+1} := \sum_{i=1}^{r+1} \omega_i$

Vector $X_m(A)$

To describe another main result, we need a notation $X_m(A)$

Definition (Vector $X_m(A)$)

Let A be an $n \times n$ Hermitian matrix. For any $1 \leq m \leq n$, let

$$X_m(A)$$

be vector containing eigenvalues of all $\binom{n}{m}$ principal submatrices of size m

Dimension:	The vector $X_m(A)$ has length $m\binom{n}{m}$
Ordering:	We assume the entries are sorted non-increasingly
Sum:	The sum of entries is $\binom{n-1}{m-1}\operatorname{tr}(A)$

Vector $X_m(A)$

Definition (Vector $X_m(A)$)

Let A be an $n \times n$ Hermitian matrix. For any $1 \leq m \leq n$, let

$$X_m(A)$$

be vector containing the eigenvalues of all principal submatrices of size m

Extreme cases

Case $m = n$: Only one principal submatrix $\implies X_n(A) = \lambda(A)$

Case $m = 1$: Submatrices are diagonal entries $\implies X_1(A) = \text{diag}(A)$

Our third main result

Using the aggregate bounds, we established a complete spectral hierarchy between eigenvalues of principal submatrices of any two sizes m and k

Theorem (H.S., M.Y.)

For any $1 \leq k \leq m \leq n$

$$\underbrace{X_m(A) \cup \cdots \cup X_m(A)}_{\binom{m-1}{k-1} \text{ times}} \succ \underbrace{X_k(A) \cup \cdots \cup X_k(A)}_{\binom{n-k}{m-k} \text{ times}}$$

This majorization generalizes Schur's theorem $(m = n, k = 1)$
This majorization implies Szasz's Inequalities by applying product

Szasz's inequalities

Let $P_m(A)$ be the product of all principal minors of size m

$$P_m(A) = \prod_{x \in X_m(A)} x$$

Apply product to the last theorem

$$(P_m(A))^{\binom{m-1}{k-1}} \leq (P_k(A))^{\binom{n-k}{m-k}}$$

Raising the inequality to the power of $1/\left[\binom{n-1}{m-1}\binom{m-1}{k-1}\right]$ yields

Szasz's inequalities

For any $n \times n$ positive semi-definite matrix A

$$P_m(A)^{1/\binom{n-1}{m-1}} \leq P_k(A)^{1/\binom{n-1}{k-1}} \quad (k \leq m)$$

Entrée

Polynomial setup

Let $\lambda_1 \geq \cdots \geq \lambda_n$ be real numbers (called poles)
Let $\omega_1, \ldots, \omega_n$ be non-negative weights such that $\sum_{i=1}^{n} \omega_i = 1$
Consider the polynomial

$$p(x) := \sum_{i=1}^{n} \omega_i \prod_{\substack{j=1 \\ j \neq i}}^{n} (x - \lambda_j) \tag{1}$$

The polynomial $p(x)$ has $n-1$ real roots $\mu_1 \geq \cdots \geq \mu_{n-1}$ which interlace with the poles

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \cdots \geq \mu_{n-1} \geq \lambda_n$$

The Rational Function Representation

If λ_i's are distinct and ω_i's are nonnegative
dividing $p(x)$ by $\prod(x-\lambda_j)$ gives the equivalent equation

$$f(x) := \sum_{i=1}^{n} \frac{\omega_i}{x-\lambda_i} = \frac{p(x)}{\prod_{j=1}^{n}(x-\lambda_j)} \tag{2}$$

μ_k's are the zeros of f if λ_i's are distinct and ω_i are nonnegative
f is strictly decreasing on $(\lambda_{i+1}, \lambda_i)$ with $f(\lambda_{i+1}+) = \infty$ and $f(\lambda_i-) = -\infty$
Thus, $f(x)$ vanishes exactly once in each such interval
These vanishing points are zeros of $p(x)$ namely μ_k's

The Main Result

Our goal is to prove the following theorem

Theorem (H.S., M.Y.)

$$\sum_{j=\ell}^{r} \lambda_{j+1} + \sum_{j=\ell}^{r} \frac{\omega_{j+1}}{L_{r+1}}(\lambda_\ell - \lambda_{j+1}) \le \sum_{j=\ell}^{r} \mu_j \le \sum_{j=\ell}^{r} \lambda_j - \sum_{j=\ell}^{r} \frac{\omega_j}{U_\ell}(\lambda_j - \lambda_{r+1})$$

where $U_\ell := \sum_{i=\ell}^{n} \omega_i$ *and* $L_{r+1} := \sum_{i=1}^{r+1} \omega_i$

Proof Strategy: Four Simplifying Assumptions

Proving it directly is cumbersome
We proceed by imposing four assumptions to simplify the problem

1. Distinct poles: $\lambda_1 > \lambda_2 > \cdots > \lambda_n$
2. Positive weights: $\omega_i > 0$ for all i
3. Rational weights: $\omega_i \in \mathbb{Q}$ for all i
4. Equal weights: $\omega_i = \frac{1}{n}$ for all i

Justifying Assumptions 1, 2 & 3

Are we allowed to make these assumptions? Yes, due to continuity

Why assumptions 1 & 2 (λ distinctness & ω positivity) make sense?
These ensure roots of p are precisely roots of f
Since roots are continuous functions of λ_i's,
results for distinct λ_i's hold for non-distinct λ_i's in the limit
Since roots are continuous functions of ω_i's,
results for positive ω_i's hold for nonnegative ω_i's in the limit

Why assumption 3 (ω rationality) makes sense?
The rational numbers $\mathbb{Q}$ are dense in $\mathbb{R}$
Any real weight ω_i is the limit of a rational sequence $\omega_i^{(k)} \to \omega_i$
Since desired inequalities are non-strict ($\leq$), they are preserved under limits

Justifying Assumption 4

Assumption 4 ($\omega_i = 1/n$) seems restrictive. Why does it make sense?

Suppose assumptions 1, 2, 3

$$\omega_i = \frac{k_i}{N}, \quad \text{where } k_i \in \mathbb{Z}^+ \text{ and } N = \sum k_i$$

Then f becomes

$$f(x) = \sum_{i=1}^{n} \frac{k_i/N}{x - \lambda_i} = \frac{1}{N} \sum_{i=1}^{n} \underbrace{\left(\frac{1}{x - \lambda_i} + \cdots + \frac{1}{x - \lambda_i} \right)}_{k_i \text{ times}}$$

We view this as a system with N poles (where pole λ_i is repeated k_i times) and equal weights $1/N$

Justifying Assumption 4

$$f(x) = \sum_{i=1}^{n} \frac{k_i/N}{x - \lambda_i} = \frac{1}{N}\sum_{i=1}^{n} \underbrace{\left(\frac{1}{x-\lambda_i} + \cdots + \frac{1}{x-\lambda_i}\right)}_{k_i \text{ times}} = \frac{1}{N}\sum_{i=1}^{n}\sum_{j=1}^{k_i}\left(\frac{1}{x-\lambda_{ij}}\right)$$

where $\lambda_{ij} = \lambda_i$ for $j = 1, 2, \ldots, k_i$
However, this violates assumption 1 (λ distinct)!
So we replace λ_{ij} by $\lambda_{ij}(\epsilon) = \lambda_i + \frac{\epsilon}{j}$

Proof (Equal Weights): Step 0 - critical points of q

When we assume $\omega_i = 1/N$, the structure becomes elegant

Let $q(x) = \prod_{j=1}^{N}(x - \lambda_j)$ be the polynomial with roots at the poles
The polynomial $p(x)$ becomes

$$p(x) = \frac{1}{N}\sum_{i=1}^{N}\prod_{j\neq i}(x - \lambda_j) = \frac{1}{N}q'(x).$$

Under assumptions 1, 2, 3, 4, the roots μ are simply the critical points (zeros of the derivative) of the polynomial $q(x)$

Proof (Equal Weights): Step 1 - Monotonicity

We now prove the bounds assuming $\omega_i = 1/n$

First, we show that roots μ move in the same direction as poles λ

Differentiating $f(\mu_k) = \sum_{i=1}^{n} \frac{1}{\mu_k - \lambda_i} = 0$ with respect to λ_j gives

$$\frac{\partial \mu_k}{\partial \lambda_j} = \frac{(\mu_k - \lambda_j)^{-2}}{\sum_{i=1}^{n} (\mu_k - \lambda_i)^{-2}} > 0$$

Every root μ_k is a strictly increasing function of every pole λ_j

Proof (Equal Weights): Step 2 - The Upper Bound

Since μ increases with λ, to maximize $\sum_\ell^r \mu_j$
we push $\lambda_1, \dots, \lambda_{\ell-1} \to +\infty$, and push $\lambda_{r+2}, \dots, \lambda_n$ up to λ_{r+1}
The roots $\mu_\ell, \dots, \mu_r$ become the critical points of

$$\tilde{q}(x) = (x - \lambda_{r+1})^{n-r} \prod_{j=\ell}^{r} (x - \lambda_j).$$

Using Vieta's formulas on $\tilde{q}'(x)$, we obtain

$$\sum_{j=\ell}^{r} \mu_j \leq \sum_{j=\ell}^{r} \mu_j = \sum_{j=\ell}^{r} \lambda_j - \frac{1}{n-\ell+1} \sum_{j=\ell}^{r} (\lambda_j - \lambda_{r+1})$$

Proof of general case: Rational approximation

Assume rational weights $w_i = k_i/N$ with integers k_i and $N = \sum k_i$

Replace each pole λ_i with a cluster of k_i distinct poles

$$\lambda_{i,s}(\epsilon) = \lambda_i + \frac{\epsilon}{s} \quad (s = 1, \ldots, k_i)$$

This creates a system with N total distinct poles and equal weights $1/N$

Apply the equal weight bound to these N poles
The count of poles from index ℓ to n is

$$N_{\text{active}} = \sum_{i=\ell}^{n} k_i = N \sum_{i=\ell}^{n} w_i = N U_\ell$$

Proof of general case: Limiting argument

The equal weight bound contains the term

$$\frac{1}{N_{\text{active}}} \sum (\text{poles} - \lambda_{r+1})$$

Summing over the cluster for λ_j contributes

$$\sum_{s=1}^{k_j} \frac{1}{NU_\ell}(\lambda_{j,s} - \lambda_{r+1}) \;\to\; \frac{k_j}{NU_\ell}(\lambda_j - \lambda_{r+1})$$

Substituting $k_j = Nw_j$ recovers exactly the term in the main theorem

$$\frac{Nw_j}{NU_\ell}(\lambda_j - \lambda_{r+1}) = \frac{w_j}{U_\ell}(\lambda_j - \lambda_{r+1})$$

Continuity extends the result to real weights

The Main Result

The derivation of lower bound is similar. Thus we proved

Theorem (H.S., M.Y.)

$$\sum_{j=\ell}^{r} \lambda_{j+1} + \sum_{j=\ell}^{r} \frac{\omega_{j+1}}{L_{r+1}}(\lambda_\ell - \lambda_{j+1}) \le \sum_{j=\ell}^{r} \mu_j \le \sum_{j=\ell}^{r} \lambda_j - \sum_{j=\ell}^{r} \frac{\omega_j}{U_\ell}(\lambda_j - \lambda_{r+1})$$

where $U_\ell := \sum_{i=\ell}^{n} \omega_i$ *and* $L_{r+1} := \sum_{i=1}^{r+1} \omega_i$

Plat Principal

Matrix Setup

1. Spectral Decomposition

Let A be an $n \times n$ Hermitian matrix: $A = V \Lambda V^* = \sum_{i=1}^{n} \lambda_i v_i v_i^*$

- $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ contains eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$.
- $V = [v_1, \ldots, v_n]$ is unitary; v_i are the orthonormal eigenvectors.

2. Principal Submatrices

Let A_k be the $(n-1) \times (n-1)$ matrix obtained by deleting the k-th row and column of A.

- We denote the eigenvalues of A_k by:

$$\mu(A_k) = (\mu_{k,1} \geq \cdots \geq \mu_{k,n-1}).$$

The Bridge to Matrices

The connection between matrix eigenvalues and polynomials is given by the following Lemma (Refer to Theorem 1 in Gau Wu 2004).

Lemma (Lemma 3.1)

Let B be the compression of A onto the subspace $u^\perp$ (where u is a unit vector). The eigenvalues of B are the roots of:

$$\sum_{i=1}^{n} |\langle u, v_i \rangle|^2 \prod_{\substack{j=1 \\ j \neq i}}^{n} (x - \lambda_j) = 0. \quad (3)$$

Key Identification (Mapping)

This is **exactly** our polynomial $p(x)$ from the Entrée!

- **Weights:** $\omega_i = |\langle u, v_i \rangle|^2$.
- **Submatrices:** For A_k, we choose $u = e_k$, so $\omega_i = |v_{i,k}|^2$.

The Translation Dictionary

- **Weights:** $w_i \longrightarrow |\langle u, v_i\rangle|^2$ (Squared projection of u onto eigenvectors).
- **Roots:** $\mu_j \longrightarrow$ Eigenvalues of the compressed matrix B.

Theorem (Theorem 3.2: Renormalized Geometric Bounds)

For any unit vector u, the eigenvalues μ_j of the compression B satisfy:

$$\sum_{j=\ell}^{r} \lambda_{j+1} + \sum_{j=\ell}^{r} \frac{|\langle u, v_{j+1}\rangle|^2}{L_{r+1}(u)}(\lambda_\ell - \lambda_{j+1}) \le \sum_{j=\ell}^{r} \mu_j$$

$$\le \sum_{j=\ell}^{r} \lambda_j - \sum_{j=\ell}^{r} \frac{|\langle u, v_j\rangle|^2}{U_\ell(u)}(\lambda_j - \lambda_{r+1}).$$

where

$$U_\ell(u) := \sum_{i=\ell}^{n} |\langle u, v_i\rangle|^2 \quad \textit{and} \quad L_{r+1}(u) := \sum_{i=1}^{r+1} |\langle u, v_i\rangle|^2.$$

The problem of bounding eigenvalues of principal submatrices corresponds to a specific choice of the projection vector: $u = e_k$.

Theorem (Theorem 3.3)

Fix an index k. The eigenvalues $\mu_{k,j}$ of the submatrix A_k satisfy:

$$\sum_{j=\ell}^{r} \lambda_{j+1} + \sum_{j=\ell}^{r} \frac{|v_{j+1,k}|^2}{L_{r+1}^k}(\lambda_\ell - \lambda_{j+1}) \leq \sum_{j=\ell}^{r} \mu_{k,j} \leq \sum_{j=\ell}^{r} \lambda_j - \sum_{j=\ell}^{r} \frac{|v_{j,k}|^2}{U_\ell^k}(\lambda_j - \lambda_{r+1}),$$

provided the renormalization factors are non-zero:

$$U_\ell^k := \sum_{i=\ell}^{n} |v_{i,k}|^2 \quad \text{and} \quad L_{r+1}^k := \sum_{i=1}^{r+1} |v_{i,k}|^2. \tag{4}$$

Significance: This formula explicitly links the spectrum of the submatrix A_k to the k-**th components** of the original eigenvectors v_i.

What happens if we sum the eigenvalues of **ALL** principal submatrices? Summing our previous result over $k = 1, \ldots, n$ yields the following bounds.

Theorem (Theorem 3.4: Aggregate Renormalized Bounds)

For any window $[\ell, r]$, *the total sum satisfies:*

$$n \sum_{j=\ell}^{r} \lambda_{j+1} + \sum_{j=\ell}^{r} (\lambda_\ell - \lambda_{j+1}) \Psi_{j+1}(r) \leq \sum_{k=1}^{n} \sum_{j=\ell}^{r} \mu_{k,j}$$

$$\leq n \sum_{j=\ell}^{r} \lambda_j - \sum_{j=\ell}^{r} (\lambda_j - \lambda_{r+1}) \Omega_j(\ell),$$

where Ω *and* Ψ *are the global renormalization factors:*

$$\Omega_j(\ell) := \sum_{k=1}^{n} \frac{|v_{j,k}|^2}{U_\ell^k} \quad \text{and} \quad \Psi_j(r) := \sum_{k=1}^{n} \frac{|v_{j,k}|^2}{L_{r+1}^k}.$$

The previous bounds depend on eigenvectors. Can we get a simpler, coordinate-free result? Yes! Since denominators are ≤ 1, we have $\Omega_j(\ell) \geq 1$ and $\Psi_j(r) \geq 1$.

Theorem (Theorem 3.5)

*Relaxing the factors yields bounds depending **only** on eigenvalues:*

$$(r-\ell+1)\lambda_\ell + (n-1)\sum_{j=\ell}^{r} \lambda_{j+1} \leq \sum_{k=1}^{n}\sum_{j=\ell}^{r} \mu_{k,j} \leq (n-1)\sum_{j=\ell}^{r} \lambda_j + (r-\ell+1)\lambda_{r+1}.$$

Connection to History

When $\ell = r = j$, this collapses exactly to Thompson's 1976 Bounds:

$$(n-1)\lambda_{j+1} + \lambda_j \leq \sum_{k=1}^{n} \mu_{k,j} \leq (n-1)\lambda_j + \lambda_{j+1}.$$

Fromage

The Spectral Hierarchy Setup

To describe the hierarchy, we need precise notation for the collections of eigenvalues.

Definition (Vectors $X_m(A)$ and $Y_{n,m}(A)$)

Let A be an $n \times n$ Hermitian matrix.

- Let $X_m(A)$ be the vector containing the eigenvalues of **ALL** $\binom{n}{m}$ principal submatrices of size $m \times m$.
- Let $Y_{n,m}(A)$ be the vector containing $\binom{n-1}{m-1}$ copies of the eigenvalues of A itself ($\lambda(A)$).

Example

- $X_n(A) = \lambda(A)$ (The full spectrum).
- $X_1(A) = \text{diag}(A)$ (The diagonal entries).

The Spectral Hierarchy Theorem

This is the "Strong Cheese" of our menu: a complete majorization ordering between submatrices of any two sizes.

Theorem (Theorem 4.1: Spectral Hierarchy)

For any integers $1 \leq k \leq m \leq n$, the eigenvalues of size m majorize the eigenvalues of size k (with appropriate repetition):

$$\underbrace{X_m(A) \cup \cdots \cup X_m(A)}_{\binom{m-1}{k-1} \text{ times}} \succ \underbrace{X_k(A) \cup \cdots \cup X_k(A)}_{\binom{n-k}{m-k} \text{ times}}. \tag{5}$$

Meaning: The spectrum becomes "more spread out" as the submatrix size k decreases.

Recovering Classics

Our hierarchy unifies several classical results.

1. Schur's Theorem ($m = n, k = 1$)

The relation simplifies to:

$$\lambda(A) \succ \operatorname{diag}(A).$$

Let $P_m(A)$ denote the product of all $m \times m$ principal minors of A.

2. Szasz's Inequalities

Since products are Schur-concave, majorization implies inequalities for products of elements. Our theorem recovers the famous Szasz's Inequalities for principal minors:

$$P_m(A)^{1/\binom{n-1}{m-1}} \leq P_k(A)^{1/\binom{n-1}{k-1}}, \quad \text{for } k \leq m.$$

Let's prepare some results to prove Theorem 4.1: Spectral Hierarchy.

Lemma (Lemma 4.3)

For any $n \geq 2$ and any $n \times n$ Hermitian matrix A:

$$Y_{n,n-1}(A) \succ X_{n-1}(A),$$

namely

$$\underbrace{\lambda(A) \cup \cdots \cup \lambda(A)}_{n-1 \text{ times}} \succ X_{n-1}(A).$$

We have

$$\sum_{i=1}^{(n-1)n} (X_{n-1})_i = \sum_{k=1}^{n} \sum_{j=1}^{n-1} \mu_{k,j} = \sum_{k=1}^{n} \operatorname{tr}(A_k) = (n-1)\operatorname{tr}(A)$$

$$= (n-1)\sum_{j=1}^{n-1} \lambda_j + (n-1)\lambda_n = \sum_{i=1}^{(n-1)n} (Y_{n,n-1})_i.$$

Proof Mechanism :

1. **Nodal Check ($p = rn$):** First, we verify the majorization inequality at integer multiples of n.

$$\sum_{i=1}^{rn}(X_{n-1})_i \le \sum_{i=1}^{rn}(Y_{n,n-1})_i.$$

   This follows from the upper bound in Theorem 3.5, applied with $\ell = 1$:

$$\sum_{i=1}^{rn}(X_{n-1})_i = \sum_{k=1}^{n}\sum_{j=1}^{r}\mu_{k,j} \le (n-1)\sum_{j=1}^{r}\lambda_j + r\lambda_{r+1} = \sum_{i=1}^{rn}(Y_{n,n-1})_i.$$

2. **Unimodality:** Between the nodes, the difference $\Delta_p = \sum_{i=1}^{p}(Y_{n,n-1})_i - \sum_{i=1}^{p}(X_{n-1})_i$ rises and falls (due to Interlacing). Since it is non-negative at the nodes, it remains non-negative everywhere.

The Inductive Step

We now establish the structural link between dimension n and $n-1$.

Lemma (Lemma 4.4: The Inductive Link)

Let $n \geq 2$. Suppose that

$$Y_{n-1,m}(B) \succ X_m(B)$$

holds for all $(n-1) \times (n-1)$ Hermitian matrices B.
Then,

$$Y_{n,m}(A) \succ X_m(A)$$

holds for all $n \times n$ Hermitian matrices A.

Significance: This allows us to use induction on the matrix size n. Since the base case $(n = 2)$ is trivial, if we prove this lemma, the theorem holds for all n.

Proof of Lemma 4.4 (Aggregation)

Step 1: Apply Hypothesis to Submatrices
Let $A_1, \ldots, A_n$ be the principal submatrices of size $(n-1)$. By assumption, for each k:

$$Y_{n-1,m}(A_k) \succ X_m(A_k).$$

By the **scaling property** of majorization, the union over all k preserves the ordering:

$$\bigcup_{k=1}^{n} Y_{n-1,m}(A_k) \succ \bigcup_{k=1}^{n} X_m(A_k). \tag{6}$$

We must now analyze the composition of the LHS and RHS vectors counting the multiplicities.

Proof of Lemma 4.4 (Counting)

Step 2: Analyze the LHS

- Each term $Y_{n-1,m}(A_k)$ contains $\binom{n-2}{m-1}$ copies of $\lambda(A_k)$.
- The union $\bigcup_{k=1}^{n} \lambda(A_k)$ is exactly the vector $X_{n-1}(A)$.
- **Total LHS:** $\binom{n-2}{m-1}$ copies of $X_{n-1}(A)$.

Step 3: Analyze the RHS

- The union collects eigenvalues of $m \times m$ submatrices found within each parent A_k.
- An $m \times m$ matrix M is contained in A_k if the index deleted to form A_k is **not** in M.
- There are $n - m$ such available indices. Thus, every M is counted exactly $n - m$ times.
- **Total RHS:** $(n - m)$ copies of $X_m(A)$.

Proof of Lemma 4.4 (Conclusion)

Substituting these counts back into the majorization relation:

$$\underbrace{X_{n-1}(A) \cup \cdots \cup X_{n-1}(A)}_{\binom{n-2}{m-1} \text{ times}} \succ \underbrace{X_m(A) \cup \cdots \cup X_m(A)}_{n-m \text{ times}}.$$

Step 4: Transitivity

From Lemma 4.3, we know $\lambda(A)$ (repeated $n-1$ times) $\succ X_{n-1}(A)$. Substituting this into the LHS:

$$\underbrace{\lambda(A) \cup \cdots \cup \lambda(A)}_{\binom{n-2}{m-1}(n-1) \text{ times}} \succ \underbrace{X_m(A) \cup \cdots \cup X_m(A)}_{n-m \text{ times}}.$$

Step 5: Cancellation

Divide the repetition counts by $(n-m)$. The new count for $\lambda(A)$ is:

$$\frac{n-1}{n-m}\binom{n-2}{m-1} = \frac{n-1}{n-m} \cdot \frac{(n-2)!}{(m-1)!(n-m-1)!} = \binom{n-1}{m-1}.$$

This matches the definition of $Y_{n,m}(A)$.

Theorem 4.5: General Majorization

We now state the general relationship between the eigenvalues of A and the submatrices of a fixed size m.

Theorem (Theorem 4.5)

For any $n \times n$ Hermitian matrix A and any submatrix size $1 \leq m \leq n$:

$$Y_{n,m}(A) \succ X_m(A).$$

Recap of Notation:

- $X_m(A)$: Vector of eigenvalues of all $\binom{n}{m}$ submatrices of size m.
- $Y_{n,m}(A)$: Vector of eigenvalues of A, repeated $\binom{n-1}{m-1}$ times.

Proof of Theorem 4.5 (Boundary Cases)

Step 1: Boundary Cases ($m = 1$ and $m = n$) The theorem holds trivially at the extremes:

- **Case $m = 1$:** $Y_{n,1}(A)$ is $\lambda(A)$. $X_1(A)$ is the vector of diagonal entries.

$$Y_{n,1}(A) \succ X_1(A) \iff \lambda(A) \succ \mathrm{diag}(A).$$

  This is precisely **Schur's Theorem**.

- **Case $m = n$:**

$$Y_{n,n}(A) = \lambda(A) = X_n(A).$$

  Equality implies majorization.

Low Dimensions ($n = 1, 2$): For $n = 1$ (where $m = 1$) and $n = 2$ (where $m = 1, 2$), the result is fully covered by these boundary cases.

Proof of Theorem 4.5 (Induction)

Step 2: Induction ($n \geq 3$) We proceed by induction on the dimension n.

Inductive Hypothesis

Assume that for all $(n-1) \times (n-1)$ Hermitian matrices B, and all $1 \leq m \leq n-1$:

$$Y_{n-1,m}(B) \succ X_m(B).$$

Applying the Lemma: By **Lemma 4.4** (The Inductive Link), if the relation holds for size $n-1$, it implies:

$$Y_{n,m}(A) \succ X_m(A)$$

for all $n \times n$ matrices A and all $1 \leq m \leq n$.

Conclusion: Since the base cases ($n = 1, 2$) are established, the theorem holds for all n by mathematical induction. □

Proof of Spectral Hierarchy (Local Application)

Step 1: Apply Theorem 4.5 Locally

Let M be a specific principal submatrix of A of size $m \times m$. We treat M as a Hermitian matrix in its own right. Apply **Theorem 4.5** to M with a smaller subsize k $(1 \leq k \leq m)$.

We have $Y_{m,k}(M) \succ X_k(M)$, which expands to:

$$\underbrace{\lambda(M) \cup \cdots \cup \lambda(M)}_{\binom{m-1}{k-1} \text{ times}} \succ X_k(M), \tag{7}$$

where $X_k(M)$ is the vector of eigenvalues of all $k \times k$ submatrices contained within M.

Proof of Spectral Hierarchy (Aggregation)

Step 2: Take the Union

We take the union of the local relationship (7) over **all** $\binom{n}{m}$ possible principal submatrices M of A.

Analyzing the Left-Hand Side (LHS):

- The union collects the vectors $\lambda(M)$ for every submatrix M.
- By definition, the collection of all $\lambda(M)$ is exactly the vector $X_m(A)$.
- Since each local term repeated $\lambda(M)$ exactly $\binom{m-1}{k-1}$ times, the aggregate LHS is:

$$\underbrace{X_m(A) \cup \cdots \cup X_m(A)}_{\binom{m-1}{k-1} \text{ times}}.$$

Proof of Spectral Hierarchy (Counting Parents)

Analyzing the Right-Hand Side (RHS): The RHS collects every $k \times k$ submatrix found inside every $m \times m$ parent M.

The Counting Argument: Let Q be a specific $k \times k$ principal submatrix of A. How many times does its spectrum appear in the union?

- Q appears whenever we select a parent M that contains Q.
- To form such a parent M, we must include the k indices of Q.
- We must choose the remaining $m - k$ indices from the available $n - k$ indices of A.

Total number of parents containing Q:

$$\text{Count} = \binom{n-k}{m-k}.$$

Proof of Spectral Hierarchy (Conclusion)

Step 3: Final Assembly
Since every distinct $k \times k$ submatrix Q is counted exactly $\binom{n-k}{m-k}$ times, the aggregate RHS is equivalent to:

$$\underbrace{X_k(A) \cup \cdots \cup X_k(A)}_{\binom{n-k}{m-k} \text{ times}}.$$

Combining the LHS and RHS yields the Theorem:

$$\underbrace{X_m(A) \cup \cdots \cup X_m(A)}_{\binom{m-1}{k-1} \text{ times}} \succ \underbrace{X_k(A) \cup \cdots \cup X_k(A)}_{\binom{n-k}{m-k} \text{ times}}.$$

□

Dessert

A Concrete Numerical Tasting

The Ingredients: A Concrete 4×4 Matrix

Instead of assuming values, let's construct a block-diagonal matrix A:

$$A = \left(\begin{array}{cc|cc} 6 & 2 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ \hline 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right)$$

1. The Eigenvalues of A:

Since A is block diagonal, its spectrum is the union of the spectra of the two 2×2 blocks.

- Top Block: Roots of $\lambda^2 - 9\lambda + 14 = 0 \implies \mathbf{7}, \mathbf{2}$.
- Bottom Block: Roots of $\lambda^2 - 4\lambda + 3 = 0 \implies \mathbf{3}, \mathbf{1}$.

$$\lambda(A) = (7, \quad 3, \quad 2, \quad 1).$$

The Submatrices (Calculated)

Deleting a row/column from a block diagonal matrix is easy. It breaks one block into a scalar, leaving the other block intact.

The 4 Principal Submatrices (3×3)**:**

- A_1 **(Delete row 1):** Leaves scalar 3 and the full Bottom Block $(3, 1)$.
  $\implies \mu(A_1) = (3, 3, 1)$.
- A_2 **(Delete row 2):** Leaves scalar 6 and the full Bottom Block $(3, 1)$.
  $\implies \mu(A_2) = (6, 3, 1)$.
- A_3 **(Delete row 3):** Leaves Top Block $(7, 2)$ and scalar 2.
  $\implies \mu(A_3) = (7, 2, 2)$.
- A_4 **(Delete row 4):** Leaves Top Block $(7, 2)$ and scalar 2.
  $\implies \mu(A_4) = (7, 2, 2)$.

Verification: Sum of all $\mu = 7 + 23 + 9 = 39$.
$(n-1)\text{tr}(A) = 3 \times (6 + 3 + 2 + 2) = 3 \times 13 = 39$. ✓

Course 1: The Power of Aggregation (Theorem 3.5)

To see the improvement over Thompson, we sum the **1st and 2nd largest eigenvalues** of all submatrices.

1. **The Actual Sum:**

$$\sum_{k=1}^{4}(\mu_{k,1} + \mu_{k,2}) = (3+3) + (6+3) + (7+2) + (7+2) = \mathbf{33}.$$

2. **The New Aggregate Bound (Thm 3.5):**

$$\mathrm{UB}_{\mathrm{New}} = (n-1)(\lambda_1 + \lambda_2) + (r - \ell + 1)\lambda_3 = 3(7+3) + 2(2) = \mathbf{34}.$$

3. **The Naive Thompson Bound (Summing separate bounds):**

$$\mathrm{UB}_{\mathrm{Naive}} = [(n-1)\lambda_1 + \lambda_2] + [(n-1)\lambda_2 + \lambda_3] = 24 + 11 = \mathbf{35}.$$

$$\mathrm{Actual}(\mathbf{33}) \leq \mathrm{New}(\mathbf{34}) < \mathrm{Naive}(\mathbf{35})$$

Conclusion: Our aggregate bound is strictly sharper than simply adding up Thompson's inequalities!

Course 2: Spectral Hierarchy (Theorem 4.1)

We check if $Y_{4,3}(A) \succ X_3(A)$.

Vector $Y_{4,3}$ (Parent)

$\lambda(A)$ repeated $\binom{3}{2} = 3$ times.

Sorted:

1. **7, 7, 7**
2. 3, 3, 3
3. 2, 2, 2
4. 1, 1, 1

Sum $=$ **39**.

Vector X_3 (Children)

Collection of all submatrix eigenvalues.

Sorted:

1. **7, 7, 6** $\leftarrow (20 < 21)$
2. 3, 3, 3
3. 2, 2, 2
4. 1, 1, 1

Sum $=$ **39**.

Majorization Check: The partial sums of Y are always $\geq$ partial sums of X.

- $k = 1$: $7 \geq 7$ ✓
- $k = 3$: $21 \geq 7 + 7 + 6 = 20$ ✓

The hierarchy holds perfectly.

This presentation is based on the following paper

Aggregate bounds on the eigenvalues of the principal submatrices of a Hermitian matrix and majorization relations

Hristo Sendov & Mengxu Yuan

DOI / Link:

http://arxiv.org/abs/2601.16320

Merci de votre attention!

Thank you for your attention!

References I

M. Adam and P. Psarrakos. On a compression of normal matrix polynomials, *Linear and Multilinear Algebra*, **54**, 85–99, (2006).

R. Bhatia. *Matrix Analysis*, Graduate Texts in Mathematics, vol. 169, Springer, New York, 1997.

J. R. Bunch, C. P. Nielsen, and D. C. Sorensen. Rank-one modification of the symmetric eigenproblem, *Numer. Math.*, **31**, 31–48, (1978).

P. B. Denton, S. J. Parke, T. Tao, and X. Zhang. Eigenvectors from eigenvalues: a survey of a basic identity in linear algebra, *Bull. Amer. Math. Soc.*, **59**, 31–58, (2022).

H.-L. Gau and P. Y. Wu. Numerical range of a normal compression, *Linear and Multilinear Algebra*, **52**, 195–201, (2004).

R. A. Horn and C. R. Johnson. *Matrix Analysis*, 2nd ed., Cambridge University Press, Cambridge, 2013.

C. R. Johnson and H. A. Robinson. Eigenvalue inequalities for principal submatrices, *Linear Algebra Appl.*, **37**, 11–22, (1981).

E.S. Katsoprinakis. Geometry of Polynomials and Majorization Theory, *Complex and harmonic analysis: Proceedings of the International Conference, May 25–27, 2007, Aristotle University of Thessaloniki*, 111-128, (2007).

A.W. Marshall, I. Olkin, and B.C. Arnold. *Inequalities: Theory of Majorization and Its Applications, 2nd ed.*, Springer, 2011.

P. Nylen, T.-Y. Tam, and F. Uhlig. On the eigenvalues of principal submatrices of normal, Hermitian and symmetric matrices, *Linear and Multilinear Algebra*, **36**, 69–78, (1993).

R. C. Thompson. Principal submatrices of normal and Hermitian matrices, *Illinois J. Math.*, **10**, 296–308, (1966).

References II

R. C. Thompson and P. McEnteggert. Principal submatrices II: The Upper and Lower Quadratic Inequalities, *Linear Algebra Appl.*, **1**, 211–243, (1968).

R. C. Thompson. Principal submatrices III: Linear Inequalities, *J. Res. Natl. Bur. Stand. Sec. B*, **72B**(1), 7–22, (1968).

R. C. Thompson. The eigenvalue spreads of a Hermitian matrix and its principal submatrices, *Linear and Multilinear Algebra*, **32**, 327–333, (1992).

R. C. Thompson and S. Therianos. Inequalities connecting the eigenvalues of a hermitian matrix with the eigenvalues of complementary principal submatrices, *Bull. Austral. Math. Soc.*, **6**, 117–132, (1972).

G. H. Golub. Some modified matrix eigenvalue problems, *SIAM Rev.*, **15**(2), 318–334, (1973).