Combinatorial Matrix Theory and Majorization

Geir Dahl
Department of Mathematics
University of Oslo

including joint work with Richard A. Brualdi

SIAM Conference on Applied Linear Algebra
Atlanta
October 2015

Majorization

Let $x, y \in \mathbb{R}^n$. $x_{[j]}$: the jth largest component in x.

Definition: x is majorized by y, denoted $x \preceq y$, if

$$\sum_{j=1}^{k} x_{[j]} \leq \sum_{j=1}^{k} y_{[j]} \quad (k < n)$$
$$\sum_{j=1}^{n} x_j = \sum_{j=1}^{n} y_j.$$

Interpretation: "x is less spread out than y": $(7,5,3) \preceq (9,4,2)$.

Generalizations: ordering matrices, measure families, group-major. etc.

- Hardy, Littlewood, Pólya, Schur, Muirhead, Dalton,...
- Arnold, Marshall and Olkin: *Inequalities: Theory of Majorization and Its Applications*, (2011) (First ed., 1979)
- Steele: *The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities*, (2004)

Basic properties

- permutation invariant: : $x \preceq Px \preceq x$ for every permutation matrix P
- Transitive, reflexive, $\preceq$ is a preorder on $\mathbb{R}^n$.
- Majorization is a partial order on the (polyhedral) cone

$$\mathcal{D}^n = \{x \in \mathbb{R}^n : x_1 \geq x_2 \geq \cdots \geq x_n\}$$

  Actually a lattice (min and max operations).
- Weak majorization: $x \preceq_w y$

Characterizations

Theorem

Let $x, y \in \mathbb{R}^n$. Equivalent:

(i) $x \preceq y$

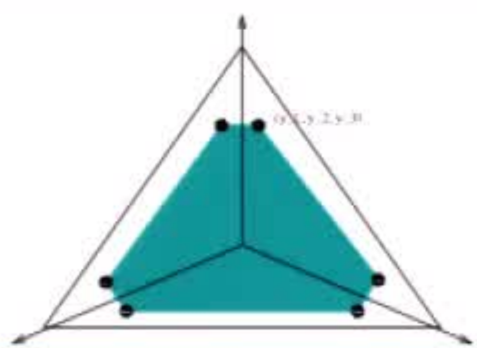
(ii) $\sum_{i=1}^{n} g(x_i) \leq \sum_{i=1}^{n} g(y_i)$ *for all convex functions* $g : \mathbb{R} \to \mathbb{R}$

(iii) $x = Ay$ *for some doubly stochastic matrix* A

(iv) $\sum_i x_i = \sum_i y_i$ *and* $\sum_i (x_i - a)^+ \leq \sum_i (y_i - a)^+$ *for all* $a \in \mathbb{R}$.

(v) $\sum_i |x_i - a| \leq \sum_i |y_i - a|$ *for all* $a \in \mathbb{R}$

(vi) x lies in the convex hull of the orbit of y under the group of permutation matrices.



Examples

- Majorization and existence results

Theorem (Schur-Horn (1923, 1954))

If $A = [a_{ij}]$ is a Hermitian matrix, with diagonal $(d_1, d_2, \ldots, d_n)$ and eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_n)$, then

$$(d_1, d_2, \ldots, d_n) \preceq (\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Conversely, if such a majorization holds in $\mathbb{R}^n$, then there exists a real symmetric matrix A with diagonal elements $d_1, d_2, \ldots, d_n$ and eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Kaftal, Weiss (2009): extension to infinite sequences, matrices.
Atiyah (1982): generalization in algebraic geometry.

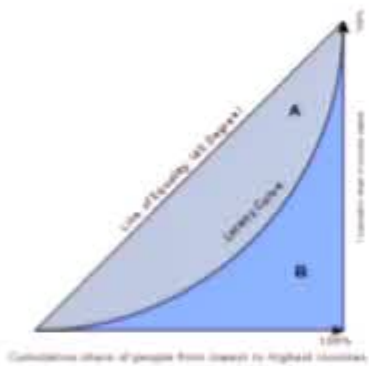
Combinatorics:

The Gale-Ryser theorem: *There exists a* $(0,1)$*-matrix with row sum vector* R *and column sum vector* S *if and only* $S \preceq R^*$.

Convexity: Doubly stochastic matrices, inequalities

Probability: Stochastic order/dominance

Economics: The Lorenz curve: income distribution, the Gini index:



Shannon information entropy: $E(p) = -\sum_i p_i \ln p_i$

Quantum physics: entanglement

- Majorization and min/max of certain symmetric functions

A function $f : \mathbb{R}^n \to \mathbb{R}$ is Schur-convex whenever $x \preceq y$ implies $f(x) \leq f(y)$. So: monotone. Then f must be symmetric.

A bounding principle: Assume f is Schur-convex on S and that $S \subseteq \mathbb{R}^n$ contains a unique *minimal* element x^1 and a unique *maximal* element x^2 in the majorization order. Then

$$\min_{x \in S} f(x) = f(x^1) \quad \text{and} \quad \max_{x \in S} f(x) = f(x^2).$$

Sometimes this gives interesting bounds; the trick is to discover an underlying majorization.

Examples: Arithmetic-geometric mean ineq., Kantorovich ineq.

And: next eigenvalues of certain Laplacian matrices ...

Laplacian energy

Discretizing the Laplace equation (a PDE)

$$\Delta u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = 0$$

gives a linear system with variables on a grid. More generally, we can consider heat flow in a graph.

Let G be a simple undirected graph with n vertices, m edges.

Laplacian matrix:

$$L(G) = D(G) - A(G)$$

where $A(G)$: adjacency matrix, $D(G)$: diagonal matrix with vertex degrees. So: $(L(G))_{ij} = -1$ when vertices $i \neq j$ are adjacent, otherwise 0, and degrees on the diagonal. Studied in Spectral graph theory.

$L(G)$: real, symmetric, positive semidefinite, with eigenvalues

$$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n = 0.$$

- Dahl, The Laplacian energy of threshold graphs and majorization (LAA, 2015). See also Helmberg, Trevisan, Threshold graphs of maximal Laplacian energy (Disc. Math, 2015). Different approaches.

Laplacian energy: $LE(G) = \sum_{i=1}^{n} |\mu_i - 2m/n|$; distance of L. eigenvalues from average degree.

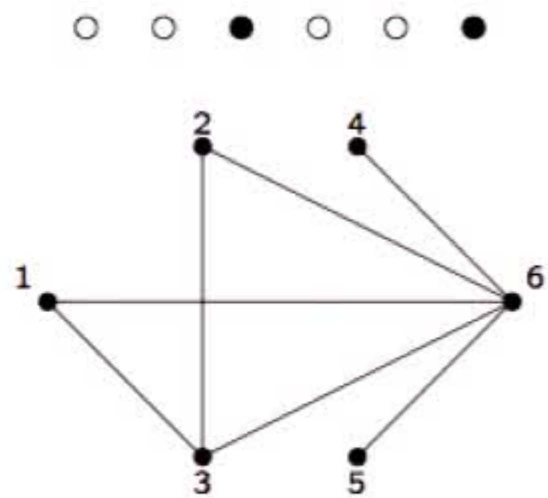
- Then $LE(G)$ is a Schur-convex function of $\mu = (\mu_1, \mu_2, \ldots, \mu_n)$, i.e., increasing w.r.t. the majorization order.
- Grone-Merris conjecture, proved by Bai: $\mu \preceq d^*$, where d^* is the conjugate sequence of degree vector d; $d_k^* = |\{i : d_i \geq k\}|$.
- So: $LE(G) \leq \sum_{i=1}^{n} |d_i^* - 2m/n|$.
- Combining Grone-Merris-Bai with Schur-Horn gives:

$$d \preceq \mu \preceq d^*.$$

Equality: $\mu = d^*$ if and only if G is a threshold graph. So: integral!!

- Goal: Find a threshold graph which maximizes the Laplacian energy
- A main message: finding underlying majorization gives nice analysis

Threshold graph: repeatedly add either an isolated vertex or a dominating vertex, which is a vertex that is connected to all vertices previously added.



Trace of G: the number of dominating vertices. Here 2.

Degree vector of a graph G with n vertices, m edges: $d = (d_1, d_2, \ldots d_n)$ is a monotone, nonnegative and integral vector with $\sum_i d_i = 2m$. Let $\kappa(d) = \max\{i : d_i \geq i\}$.

From graph theory (Ruch and Gutman): d is the degree sequence of a graph if and only if

$$\sum_{i=1}^{k} d_i \leq \sum_{i=1}^{k} (d_i^* - 1) \quad (k \leq \kappa(d)).$$

Threshold graphs: Equality here, so $d_i = d_i^* - 1$ $(i \leq \kappa(d))$.
This makes is easy to construct all threshold graphs, and see Laplacian eigenvalues.

Example: Let $n = 6$ and $m = 7$.

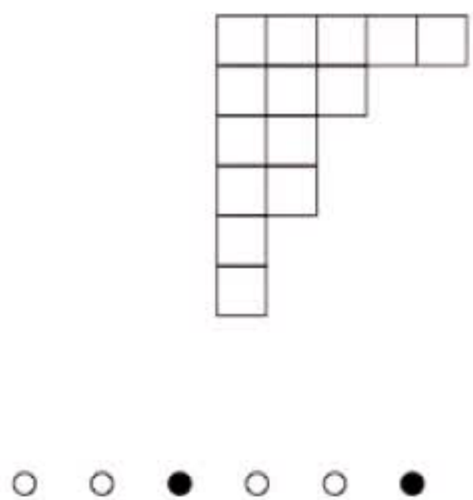
*	*	.	.	.
*	*	.		
*	*			
.	.			
.				
.				

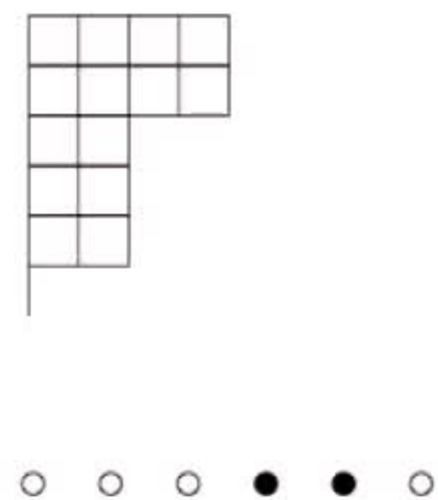
Here the Laplacian eigenvalue vector is $\mu = (6, 4, 2, 1, 1, 0)$.

We fix n, m and $\kappa(G)$. Remember: $LE(G) = \sum_{i=1}^{n} |d_i^* - 2m/n|$.

Lemma: $LE(G) = 2\sum_{i=1}^{n}(d_i^* - \alpha)^+$ where $\alpha = 2m/n$.

Example: $n = 6$, $m = 7$ and $\alpha = 2m/n = 7/3$. Move the blocks!





$\mu(G) = (6, 4, 2, 1, 1, 0)$
$LE(G) = 32/3$

$\mu(G') = (5, 5, 2, 2, 0, 0)$
$LE(G') = 32/3$

Connection to majorization: integer partitions.

Let $p, q > 0$ and $N \geq 0$ be three integers with $N \leq pq$.
Let $\mathcal{P}_{p,q}^N$ be the set of all integer partitions of N where each part x_i is bounded by q, i.e., nonincreasing integral vectors $x = (x_1, x_2, \ldots, x_p) \in \mathbb{R}^p$ satisfying

$$\sum_{i=1}^{p} x_i = N \text{ and } 0 \leq x_i \leq q \;\; (i \leq p).$$

Majorization on $\mathcal{P}_{p,q}^N$; poset: $x \preceq y$ means that $\sum_{i=1}^k x_i \leq \sum_{i=1}^k y_i$ for $k \leq p$. Define:

- $\hat{x} = (q, q, \ldots, q, r, 0, \ldots, 0)$ where the number of q's is $\lfloor N/q \rfloor$ and $r = N - \lfloor N/q \rfloor q$.
- $\tilde{x} = (v+1, \ldots, v+1, v, \ldots, v)$, where $v = \lfloor N/p \rfloor$ and the number of components being $v+1$ is $N - pv$.

Example: $p = q = 4$ and $N = 6$. Then $\hat{x} = (4,2,0,0)$ and $\tilde{x} = (2,2,1,1)$ and their Ferrers diagrams are

Lemma

$\hat{x}$ is the unique maximal element and $\tilde{x}$ is the unique minimal element in the poset $(\mathcal{P}^N_{p,q}, \preceq)$. Therefore,

$$\tilde{x} \preceq x \preceq \hat{x} \quad \textit{for all } x \in \mathcal{P}^N_{p,q}.$$

Minimal and maximal threshold degree vectors. Let $n = 7$, $m = 8$. So $2 \leq k \leq 3$.

$$\hat{F}^{(2)} = \left[\begin{array}{cc|ccccc} 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right], \qquad \tilde{F}^{(2)} = \left[\begin{array}{cc|ccccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

and

$$\hat{F}^{(3)} = \left[\begin{array}{ccc|cccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right], \qquad \tilde{F}^{(3)} = \left[\begin{array}{ccc|cccc} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Moreover

$$\begin{array}{ll} \hat{d}^{(2)} = (6,3,2,2,1,1,1), & \tilde{d}^{(2)} = (5,4,2,2,2,1,0), \\ \hat{d}^{(3)} = (5,3,3,3,1,1,0), & \tilde{d}^{(3)} = (4,4,3,3,2,0,0), \\ \hline \hat{\mu}^{(2)} = (7,4,2,1,1,1,0), & \tilde{\mu}^{(2)} = (6,5,2,2,1,0,0), \\ \hat{\mu}^{(3)} = (6,4,4,1,1,0,0), & \tilde{\mu}^{(3)} = (5,5,4,2,0,0,0). \end{array}$$

Inequalities for Laplacian energy for minimal/maximal threshold graphs:

$$L(\hat{d}^{(k_1)}) \geq L(\hat{d}^{(k_1+1)}) \geq \cdots \geq L(\hat{d}^{(\lfloor\alpha\rfloor)}) \leq L(\hat{d}^{(\lceil\alpha\rceil)}) \leq \cdots \leq L(\hat{d}^{(k_2)}),$$
$$L(\tilde{d}^{(k_1)}) \geq L(\tilde{d}^{(k_1+1)}) \geq \cdots \geq L(\tilde{d}^{(\lfloor\alpha\rfloor)}) \leq L(\tilde{d}^{(\lceil\alpha\rceil)}) \leq \cdots \leq L(\tilde{d}^{(k_2)}),$$
$$L(\hat{d}^{(k)}) \geq L(\tilde{d}^{(k)}) \qquad (k_1 \leq k \leq \lfloor\alpha\rfloor),$$
$$L(\hat{d}^{(k)}) \leq L(\tilde{d}^{(k)}) \qquad (\lceil\alpha\rceil \leq k \leq k_2).$$

Theorem

Let n and m be positive integers. Then

$$\Delta^{LE}_{n,m} = 2\max\{L(\hat{d}^{(k_1)}), L(\tilde{d}^{(k_2)})\},$$

so the Laplacian energy in $\mathcal{T}_{n,m}$ is maximized by one of the two (n,m)-extreme threshold degree vectors $\hat{d}^{(k_1)}$ and $\tilde{d}^{(k_2)}$.

Extensions: minimize, or min/max among connected threshold graphs

Qualitative matrix theory deals with matrix properties that only depend on the signs on the entries of the matrix. Motivation from economic models (P. Samuelson). See Brualdi and Shader: *Matrices of Sign-Solvable Linear Systems*, (1995).

The qualitative class of a matrix A consists of those matrices with same signs on its entries as A.

- Brualdi and Dahl, Strict sign-central matrices, SIAM Matrix Analysis Appl. (2015).

Let A be a real matrix. Define:

- A is strict central: A has a positive vector in its null space.
- A is strict sign-central (SSC-matrix): each matrix in the qualitative class of A is strict central.

Related work:

- Ando and Brualdi, Sign-central matrices, (1994).
- Lee and Shader, Sign-consistency and solvability of constrained linear systems, (1998).

A matrix A is central whenever it has a nonzero nonnegative vector in its null space.

Geometrically: the origin is in the convex hull of the columns of A.

Sign-central: each matrix in the qualitative class of A is central.

Motivation: discrete financial market

- a matrix $P = [p_{ij}]$: rows correspond to scenarios, columns to assets.
- p_{ij}: rel. change in the value of asset j for scenario i (one time step).
- portfolio: a vector $x \in \mathbb{R}^n$ where x_j is the quantity of asset j an investor holds from time t_0 to t_1.
- Px the payoff of x for each of the scenarios.
- arbitrage: Px is nonnegative, but nonzero.
- The fundamental theorem of asset pricing/mathematical finance: *there is no arbitrage if and only if there is a probability measure on the set of scenarios which makes each asset price process a martingale.* This means that there is a positive vector in the null space of P^T, i.e., P^T is strict central.

So:

- P^T is a strict central matrix iff the market is arbitrage-free
- P^T is strict sign-central iff all markets in the qualitative class of P are arbitrage-free.

This is a robustness question, motived by uncertainty in the data p_{ij}.

Example: An SSC matrix:

$$F_m = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix}.$$

Farkas' lemma/duality gives:

Theorem

Let $A \in M_{m,n}$. Then the following statements are equivalent:

(i) A is a strict central matrix.

(ii) The only nonnegative vector in the row space of A is the zero vector.

A diagonal matrix D is called a strict signing if its diagonal entries are ± 1.

Theorem (Ando and Brualdi (94))

For every $m \times n$ $(0, \pm 1)$-matrix A, the following are equivalent:

(i) A is a sign-central matrix.

(ii) For every strict signing D of order m, the matrix DA contains a nonnegative column.

The next theorem contains a characterization of the SSC property, see also Lee and Shader.

A diagonal matrix D is called a signing if its diagonal entries are $0, -1, 1$.

Theorem

Let A be an $m \times n$ $(0, \pm 1)$-matrix with no zero rows or columns. Then the following are equivalent:

(i) A is an SSC-matrix.

(ii) For every signing $D \neq O$, the matrix DA contains a nonzero nonnegative column.

This may be interpreted in the model of a financial market: For each simple nonzero portfolio x there is a scenario $i \leq m$ such that its payoff vector is nonpositive and nonzero.

The next result gives an upper bound on the number of columns of a minimal SSC matrix, i.e., an SSC matrix where no column can be deleted without destroying the SSC property.

Theorem

Let A be an $m \times n$ $(0, \pm 1)$-matrix which is minimal SSC.
Then $n \leq 2^m$.

If $n = 2^m$, then A equals (up to column permutations) the matrix E_m.

Majorization for partially ordered sets

Brualdi and Dahl, Majorization for partially ordered sets, *Discrete Math.*, 2013

Majorization extensions:

- Choquet ordering: μ, ν probability measures on a topological vector space X: ν is a dilation of μ if $\int \phi \, d\mu \leq \int \phi \, d\nu$ for all cont. convex functions on X. Phelps (1966), Meyer (1966), Alfsen (1971, 2008)
- Majorization for measure families: Blackwell (1951), Karlin, Rinott (1983), Torgersen (1968, 1985, 1991)
- Majorization induced by convex cones and groups: Marshall, Walkup, Wets (1977), Niezgoda (1998, 2007), Eaton (1984), Eaton, Perlman (1977), Lewis (1996), Tam (2000)
- Matrix majorization, polytopes etc: Hwang, Pyo (2001), Brualdi (1984), Brualdi, Hwang (1996), Dahl (1999, 2001, 2008)