Schur polynomials, entrywise positivity preservers, and weak majorization

Apoorva Khare
(Indian Institute of Science)

Terence Tao
(University of California at Los Angeles)

Abstract

- Which functions when applied *entrywise*, preserve positive semidefiniteness?
- In all dimensions, results of Schur and others show: these are power series with non-negative coefficients.
- For matrices of a fixed dimension, *no other polynomials/power series* were known.
- We provide (i) examples of such polynomials, (ii) characterization results for 'fewnomials', and (iii) a complete solution to *which* coefficients can be negative.
- The proofs crucially rely on Schur polynomials.
- Also extend a conjecture of Cuttler–Greene–Skandera.

Schur polynomials

Definition 1 (Littlewood). Given $N \in \mathbb{N}$ and decreasing integers
$$n_{N-1} > n_{N-2} > \cdots > n_0 \geqslant 0,$$
let $\mathbf{n} := (n_{N-1}, \dots, n_0)$. The Schur polynomial $s_{\mathbf{n}}(u_1, \dots, u_N)$ is the sum of weights of all column-strict Young tableaux, with shape $(n_{N-1} - (N-1), \dots, n_0 - 0)$ and alphabet $u_1, \dots, u_N$.

Example 2. Suppose $N = 3$ and $\mathbf{n} = (4, 2, 0)$. The tableaux are:

[3 3 / 2] [3 3 / 1] [3 2 / 2] [3 2 / 1] [3 1 / 2] [3 1 / 1] [2 2 / 1] [2 1 / 1]

$$\begin{aligned} &s_{(4,2,0)}(u_1, u_2, u_3) \\ &= u_3^2 u_2 + u_3^2 u_1 + u_3 u_2^2 + 2u_3 u_2 u_1 + u_3 u_1^2 + u_2^2 u_1 + u_2 u_1^2 \\ &= (u_1 + u_2)(u_2 + u_3)(u_3 + u_1). \end{aligned}$$

Schur polynomials are homogeneous and symmetric.
(Characters of irreducible polynomial representations of $GL_n(\mathbb{C})$.)

Definition 3 (Cauchy). Weyl Character Formula: $s_{\mathbf{n}}(\mathbf{u}) = \frac{\det(u_j^{n_{k-1}})}{\det(u_j^{k-1})}$.

Ratios of Schur polynomials: Monotonicity, via Schur positivity

Suppose $\mathbf{0} \leqslant \mathbf{n} \leqslant \mathbf{m}$ coordinatewise. How does the ratio of these Schur polynomials behave on the positive orthant?
$$f(\mathbf{u}) := \frac{s_{\mathbf{m}}(\mathbf{u})}{s_{\mathbf{n}}(\mathbf{u})}, \qquad \mathbf{u} \in (0, \infty)^N. \tag{4}$$

Example 5. Suppose $\mathbf{m} = (4, 2, 0)$ as above, and $\mathbf{n} = (3, 2, 0)$. Then:
$$f(u_1, u_2, u_3) = \frac{(u_1 + u_2)(u_2 + u_3)(u_3 + u_1)}{u_1 u_2 + u_2 u_3 + u_3 u_1}, \qquad u_1, u_2, u_3 > 0.$$

Note: both numerator and denominator are monomial-positive, hence non-decreasing in each coordinate.

Claim: *f is also non-decreasing in each coordinate.*

(Why?) Applying the quotient rule of differentiation to f,
$$s_{\mathbf{n}}(\mathbf{u}) \partial_{u_3} s_{\mathbf{m}}(\mathbf{u}) - s_{\mathbf{m}}(\mathbf{u}) \partial_{u_3} s_{\mathbf{n}}(\mathbf{u}) = (u_1 + u_2)(u_1 u_2 + 2u_1 u_2 + u_2 u_3) u_3,$$
and this is monomial-positive. In fact, more holds: if we write this as $\sum_{j \geqslant 0} p_j(u_1, u_2) u_3^j$ then each p_j is Schur-positive, i.e. a sum of Schur polynomials: $p_0(u_1, u_2) = 0$, while

$$p_1(u_1, u_2) = 2u_1 u_2^2 + 2u_1^2 u_2 = 2\,[2\ 2 / 1] + 2\,[2\ 1 / 1] = 2s_{(3,1)}(u_1, u_2),$$
$$p_2(u_1, u_2) = (u_1 + u_2)^2 = [2\ 2] + [2\ 1] + [1\ 1] + [2 / 1] = s_{(3,0)}(u_1, u_2) + s_{(2,1)}(u_1, u_2).$$

This happens in greater generality:

Theorem 6 (Monotonicity of Schur polynomial ratios, [1])

Suppose $0 \leqslant n_0 < \cdots < n_{N-1}$ and $0 \leqslant m_0 < \cdots < m_{N-1}$ are integers satisfying: $n_j \leqslant m_j \ \forall j$. Then the function $f(\mathbf{u}) := \frac{s_{\mathbf{m}}(\mathbf{u})}{s_{\mathbf{n}}(\mathbf{u})}$ as in (4) is non-decreasing in each u_j. More strongly: writing
$$s_{\mathbf{n}}(\mathbf{u}) \cdot \partial_{u_N} s_{\mathbf{m}}(\mathbf{u}) - s_{\mathbf{m}}(\mathbf{u}) \cdot \partial_{u_N} s_{\mathbf{n}}(\mathbf{u})$$
as a polynomial in u_N, the coefficient of each monomial u_N^j is a Schur-positive polynomial in $(u_1, u_2, \dots, u_{N-1})$.

The proof ultimately reduces to showing the Schur positivity of (i) products of skew-Schur polynomials (which follows from the Littlewood–Richardson rule); and of (ii) the expressions
$$s_{(\lambda \vee \nu)/(\mu \vee \rho)} s_{(\lambda \wedge \nu)/(\mu \wedge \rho)} - s_{\lambda/\mu} s_{\nu/\rho},$$
where λ/μ and ν/ρ are arbitrary skew shapes, and one defines
$$\lambda \vee \nu := (\max(\lambda_1, \mu_1), \max(\lambda_2, \mu_2), \dots), \ \lambda \wedge \nu := (\min(\lambda_1, \mu_1), \min(\lambda_2, \mu_2), \dots).$$

The Schur positivity of these expressions was proved by Lam, Postnikov, and Pylyavskyy in [*Amer. J. Math.* 2007].

Entrywise p.s.d. preservers: history

Let $\mathbb{P}_N(I) :=$ symmetric $N \times N$ positive semidefinite matrices with entries in $I \subset \mathbb{R}$. A function $f : I \to \mathbb{R}$ acts entrywise on $\mathbb{P}_N(I)$ via: $f[A] := (f(a_{jk}))$.

Problem

Given a function $f : \mathbb{R} \to \mathbb{R}$, when is it true that
$$f[A] = (f(a_{jk})) \in \mathbb{P}_N(\mathbb{R}) \text{ for all } A \in \mathbb{P}_N(\mathbb{R})?$$

(Note: $f(x) = 1, x$ work.) This problem has a long history, starting with Schur. The Schur product of matrices A, B is $A \circ B := (a_{jk} b_{jk})$.

Theorem 7 (Schur, *Crelle*, 1911). $A, B \in \mathbb{P}_N(\mathbb{R}) \Rightarrow A \circ B \in \mathbb{P}_N(\mathbb{R})$.

As a consequence of the Schur product theorem:

- $f(x) = x^2, x^3, \dots$ preserve positivity on $\mathbb{P}_N(\mathbb{R})$ for all N.
- Hence, if $c_k \geqslant 0 \ \forall k$ and $f(x) = \sum_{k=0}^\infty c_k x^k$ is convergent, then $f[-]$ preserves positivity in all dimensions N.

Question 8 (Pólya–Szegö, 1925). Any other preservers (for all N)?
(A "No" would show a "converse" to the Schur product theorem.)

Theorem 9 (Schoenberg, *Duke* 1942; see also Rudin, *Duke* 1959)

Let $f : (-1, 1) \to \mathbb{R}$. The following are equivalent:

1. The entrywise map preserves positivity, $f[-] : \mathbb{P}_N((-1, 1)) \to \mathbb{P}_N(\mathbb{R})$ for all $N \geqslant 1$.
2. The function f is analytic on $(-1, 1)$ and has non-negative Maclaurin coefficients.

This characterizes positivity preservers for matrices of all dimensions.

Challenging refinement: Classify entrywise maps preserving positivity in fixed dimension.

- Complete characterization **open to date** for each $N \geqslant 3$.
- Studied by Loewer, Horn, FitzGerald (1960s, 1970s) in connection with Bieberbach conjecture.
- Recently by Bhatia, Elsner, Fallat, Hiai, Jain, Johnson, Sokal, …
- Leads to a new graph invariant. [*JCT-A* 2016; *FPSAC* 2017]
- Strong modern motivations, from high-dimensional covariance estimation in applied fields.

Entrywise polynomial preservers

Today: only work with polynomial preservers, in fixed dimension N.

- If all coefficients of such a preserver f are non-negative, then $f[-] : \mathbb{P}_n(\mathbb{R}) \to \mathbb{P}_n(\mathbb{R})$ for all dimensions n.
- For fixed N, the test set is reduced, so more polynomials should work – entrywise polynomial preservers with negative coefficients.

But, *no examples known on* $\mathbb{P}_N(\mathbb{R})$! (I.e., on 'covariance' matrices.)

- If we restrict to bounded domains $[0, \rho]$ for $\rho > 0$ ('correlation' matrices), the only known necessary condition (essentially due to Horn) is that the first N nonzero coefficients of any entrywise polynomial preserver must be positive. E.g., if
$$f(x) = c_0 + c_2 x^2 + c_5 x^5 + c' x^M \quad (M > 5)$$
entrywise preserves positivity on $\mathbb{P}_3([0, \rho])$, then $c_0, c_2, c_5 \geqslant 0$; and if $c' < 0$ then $c_0, c_2, c_5 > 0$.

Now can c' be negative? What is a sharp bound?
(Open to date.)

Main result

Theorem 10 ([1] 2017)

Fix $N \geqslant 1$ and integers $0 \leqslant n_0 < \cdots < n_{N-1} < M$, and let
$$f(x) := c_{n_0} x^{n_0} + \cdots + c_{n_{N-1}} x^{n_{N-1}} + c' x^M, \tag{11}$$
where $c_{n_j} > 0$. Let $0 < \rho < \infty$. Then the following are equivalent.

1. $f[-]$ preserves positivity on $\mathbb{P}_N((0, \rho))$.
2. $\det f[A] \geqslant 0$ for all real rank-one matrices in $\mathbb{P}_N((0, \rho))$.
3. Either $c_0, \dots, c_{N-1}, c' \geqslant 0$, or $c_0, \dots, c_{N-1} > 0 > c'$ and
$$\frac{1}{|c'|} \geqslant \sum_{j=0}^{N-1} \frac{s_{\mathbf{n}_j}(1, \dots, 1)^2}{s_{\mathbf{n}}(1, \dots, 1)^2} \frac{\rho^{M - n_j}}{c_{n_j}}, \tag{12}$$
where $\mathbf{n}_j := (M, n_{N-1}, \dots, n_{j+1}, \widehat{n_j}, n_{j-1}, \dots, n_0)$ for all j.

Some consequences of Theorem 10:

- First examples of polynomial preservers with negative coefficients.
- More strongly, we can use this to characterize the sign patterns of power series preservers on $\mathbb{P}_N$.
- Provides the first construction of polynomials that preserve positivity on $\mathbb{P}_N$, but not on $\mathbb{P}_{N+1}$. Thus Horn's result is sharp.

From determinants to Schur polynomials

Sketch of proof of Theorem 10: (1) $\implies$ (2): Immediate.

(3) $\implies$ (1): Done using Theorem 6; see [1] for details.

(2) $\implies$ (3): We will only show how Schur polynomials and the threshold (12) arise out of Theorem 10(2), assuming that $c_{n_j} > 0 > c'$. First note that for any vector $\mathbf{u} \in \mathbb{R}^N$,
$$f[\mathbf{u}\mathbf{u}^T] = \sum_{j=0}^{N-1} c_{n_j} \mathbf{u}^{\circ n_j} (\mathbf{u}^{\circ n_j})^T + c' \mathbf{u}^{\circ M} (\mathbf{u}^{\circ M})^T$$
$$= (\mathbf{u}^{\circ n_0} | \dots | \mathbf{u}^{\circ n_{N-1}} | \mathbf{u}^{\circ M}) \begin{pmatrix} c_{n_0} & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & c_{n_{N-1}} & 0 \\ 0 & \cdots & 0 & c' \end{pmatrix} (\mathbf{u}^{\circ n_0} | \dots | \mathbf{u}^{\circ n_{N-1}} | \mathbf{u}^{\circ M})^T.$$

Let $V(\mathbf{u}) := \prod_{1 \leqslant j < k \leqslant N} (u_j - u_k)$ be the "Vandermonde determinant for $\mathbf{u}$". Then by the Cauchy–Binet formula and Definition 3,
$$\det f[\mathbf{u}\mathbf{u}^T] = V(\mathbf{u})^2 \prod_{j=0}^{N-1} c_{n_j} \left(c' \sum_{j=0}^{N-1} \frac{s_{\mathbf{n}_j}(\mathbf{u})^2}{c_{n_j}} + s_{\mathbf{n}}(\mathbf{u})^2 \right) \geqslant 0. \tag{13}$$

Now suppose the coordinates of $\mathbf{u}$ are distinct and in $(0, \sqrt{\rho})$. Solving for c' yields
$$\frac{1}{|c'|} \geqslant \sum_{j=0}^{N-1} \frac{s_{\mathbf{n}_j}(\mathbf{u})^2}{c_{n_j} s_{\mathbf{n}}(\mathbf{u})^2}.$$

Finally, if Theorem 10(2) holds, then $1/|c'|$ must exceed the right-hand side for all $\mathbf{u} \in (0, \sqrt{\rho})^N$ with distinct coordinates, hence for all $\mathbf{u} \in (0, \sqrt{\rho})^N$. By continuity and Theorem 6,
$$\frac{1}{|c'|} \geqslant \sum_{j=0}^{N-1} \sup_{\mathbf{u} \in (0, \sqrt{\rho})^N} \frac{s_{\mathbf{n}_j}(\mathbf{u})^2}{c_{n_j} s_{\mathbf{n}}(\mathbf{u})^2} = \sum_{j=0}^{N-1} \frac{s_{\mathbf{n}_j}(\mathbf{u}_1)^2}{c_{n_j} s_{\mathbf{n}}(\mathbf{u}_1)^2},$$
where $\mathbf{u}_1 := \sqrt{\rho}(1, \dots, 1)^T$. But this equals the threshold in (12) by the principal specialization of the (type A) Weyl Character Formula. □

Cuttler–Greene–Skandera conjecture, and weak majorization

Definition 14. Given integers $0 \leqslant n_0 \leqslant \cdots \leqslant n_{N-1}$ and $0 \leqslant m_0 \leqslant \cdots \leqslant m_{N-1}$, the vector $\mathbf{m} := (m_{N-1}, \dots, m_0)^T$ weakly majorizes $\mathbf{n} := (n_{N-1}, \dots, n_0)^T$ if
$$\sum_{j=k}^{N-1} m_j \geqslant \sum_{j=k}^{N-1} n_j, \ \forall 0 < k < N, \qquad \sum_{j=0}^{N-1} m_j \geqslant \sum_{j=0}^{N-1} n_j.$$

If moreover the final inequality is an equality, we say $\mathbf{m}$ *majorizes* $\mathbf{n}$.

A conjecture of Cuttler–Greene–Skandera [*Eur. J. Comb.* 2011] says that $\mathbf{m}$ majorizes $\mathbf{n}$ if and only if the following inequality holds:
$$\frac{s_{\mathbf{m}}(\mathbf{u})}{s_{\mathbf{n}}(\mathbf{u})} \geqslant \frac{s_{\mathbf{m}}(1, \dots, 1)}{s_{\mathbf{n}}(1, \dots, 1)}, \qquad \forall \mathbf{u} \in (0, \infty)^N.$$

The conjecture was recently proved by Sra [*Eur. J. Comb.* 2016], and Ait-Haddou and Mazure [*Found. Comput. Math.* 2018].

Compare this with Theorem 6, which says that if $\mathbf{m} \geqslant \mathbf{n}$ coordinatewise, then
$$\frac{s_{\mathbf{m}}(\mathbf{u})}{s_{\mathbf{n}}(\mathbf{u})} \geqslant \frac{s_{\mathbf{m}}(1, \dots, 1)}{s_{\mathbf{n}}(1, \dots, 1)}, \qquad \forall \mathbf{u} \in [1, \infty)^N.$$

Question 15. How can these two inequalities be reconciled?

The above three papers all assume $\sum_j m_j = \sum_j n_j$. Replace by an inequality $\rightsquigarrow$ novel characterization of weak majorization:

Theorem 16 ([1], 2017)

Given vectors $\mathbf{m}, \mathbf{n} \in \mathbb{Z}_{\geqslant 0}^N$ with strictly decreasing integer coordinates, we have
$$\frac{s_{\mathbf{m}}(\mathbf{u})}{s_{\mathbf{n}}(\mathbf{u})} \geqslant \frac{s_{\mathbf{m}}(1, \dots, 1)}{s_{\mathbf{n}}(1, \dots, 1)}, \qquad \forall \mathbf{u} \in [1, \infty)^N$$
if and only if $\mathbf{m}$ weakly majorizes $\mathbf{n}$.
(In fact, this extends to *real* tuples $\mathbf{m}, \mathbf{n} \in (0, \infty)^N$.)

Acknowledgments

- SERB (Govt. of India): Ramanujan Fellowship
- Infosys Foundation (India): Young Investigator Award
- Simons Foundation (USA): Simons Investigator Award
- NSF (USA): Grant DMS-1266164

Reference

[1] A. Khare and T. Tao. On the sign patterns of entrywise positivity preservers in fixed dimension. *Preprint*, arXiv:1708.05197, 2017.