Preservers for classes of positive matrices

Alexander Belton

Department of Mathematics and Statistics
Lancaster University, United Kingdom

a.belton@lancaster.ac.uk

18th Workshop: Noncommutative Probability, Operator Algebras, Random Matrices and Related Topics, with Applications

Będlewo

20th July 2018

Structure of the talk

Joint work with Dominique Guillot (Delaware), Apoorva Khare (IISc Bangalore) and Mihai Putinar (UCSB and Newcastle)

1. Two products on matrices
2. Positive definiteness
3. Theorems of Schur, Schoenberg and Horn
4. Positivity preservation
    1. For fixed dimension
    2. For moment matrices
    3. For totally positive matrices

Two matrix products

Notation

The set of $n \times n$ matrices with entries in a set $K \subseteq \mathbb{C}$ is denoted $M_n(K)$.

Products

The vector space $M_n(\mathbb{C})$ is an associative algebra for at least two different products: if $A = (a_{ij})$ and $B = (b_{ij})$ then

$$\begin{aligned} (AB)_{ij} &:= \sum_{k=1}^{n} a_{ik} b_{kj} \quad \text{(standard)} \\ \text{and} \quad (A \circ B)_{ij} &:= a_{ij} b_{ij} \quad \text{(Hadamard).} \end{aligned}$$

Positive definiteness

Definition

A matrix $A \in M_n(\mathbb{C})$ is *positive semidefinite* if

$$\mathbf{x}^* A \mathbf{x} = \sum_{i,j=1}^{n} \overline{x_i} a_{ij} x_j \geqslant 0 \qquad \text{for all } \mathbf{x} \in \mathbb{C}^n.$$

A matrix $A \in M_n(\mathbb{C})$ is *positive definite* if

$$\mathbf{x}^* A \mathbf{x} = \sum_{i,j=1}^{n} \overline{x_i} a_{ij} x_j > 0 \qquad \text{for all } \mathbf{x} \in \mathbb{C}^n \setminus \{0\}.$$

Remark

The subset of $M_n(K)$ consisting of positive semidefinite matrices is denoted $M_n(K)_+$.The set $M_n(\mathbb{C})_+$ is a *cone*: closed under sums and under multiplication by elements of $\mathbb{R}_+$.

Some consequences

Symmetry

If $A \in M_n(K)_+$ then $A^T \in M_n(K)_+$: note that

$$0 \leqslant (\mathbf{x}^* A \mathbf{x})^T = \mathbf{y}^* A^T \mathbf{y} \qquad \text{for all } \mathbf{y} = \overline{\mathbf{x}} \in \mathbb{C}^n.$$

Hermitianity

If $A \in M_n(\mathbb{C})_+$ then $A = A^*$: note that

$$\mathbf{x}^* A^* \mathbf{x} = (\mathbf{x}^* A \mathbf{x})^* = \mathbf{x}^* A \mathbf{x} \quad (\mathbf{x} \in \mathbb{C}^n) \quad \Longrightarrow \quad A = A^* \quad \text{(polarisation)}.$$

Non-negative eigenvalues

If $A \in M_n(\mathbb{C})_+ = U^* \operatorname{diag}(\lambda_1, \ldots, \lambda_n) U$, where $U \in M_n(\mathbb{C})$ is unitary, then $\lambda_i = (U^* \mathbf{e}_i)^* A (U^* \mathbf{e}_i) \geqslant 0$.

Conversely, if $A \in M_n(\mathbb{C})$ is Hermitian and has non-negative eigenvalues then $A = B^* B \in M_n(\mathbb{C})_+$, where $B = A^{1/2} = U^* \operatorname{diag}\big(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}\big) U$.

The Schur product theorem

Theorem 1 (Schur, 1911)

If $A, B \in M_n(\mathbb{C})_+$ *then* $A \circ B \in M_n(\mathbb{C})_+$.

Proof.

$$\begin{aligned} \mathbf{x}^*(A \circ B)\mathbf{x} &= \operatorname{tr}(\operatorname{diag}(\mathbf{x})^* B \operatorname{diag}(\mathbf{x}) A^T) \\ &= \operatorname{tr}\big((A^T)^{1/2} \operatorname{diag}(\mathbf{x})^* B \operatorname{diag}(\mathbf{x}) (A^T)^{1/2}\big). \end{aligned}$$

□

A question of Pólya and Szegö

Corollary 2

If $f(z) = \sum_{n=0}^{\infty} a_n z^n$ *is analytic on* K *and* $a_n \geqslant 0$ *for all* $n \geqslant 0$ *then*

$$f[A] := \big(f(a_{ij})\big) \in M_n(\mathbb{C})_+ \qquad \textit{for all } A = (a_{ij}) \in M_n(K)_+ \textit{ and all } n \geqslant 1.$$

Observation

The functions in the corollary preserve positivity on $M_n(K)_+$ regardless of the dimension n.

Question (Pólya–Szegö, 1925)

Are there any other functions with this property?

Schoenberg's theorem

Theorem 3 (Schoenberg, 1942)

If $f : [-1,1] \to \mathbb{R}$ *is*

(i) *continuous and*

(ii) *such that* $f[A] \in M_n(\mathbb{R})_+$ *for all* $A \in M_n\big([-1,1]\big)_+$ *and all* $n \geqslant 1$

then f *is absolutely monotonic:*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \qquad \text{for all } x \in [-1,1], \quad \text{where } a_n \geqslant 0 \text{ for all } n \geqslant 0.$$

Proof (Christensen–Ressel, 1978).

$$\{f : [-1,1] \to \mathbb{R} \mid f(1) = 1,\ f[A] \in M_n(\mathbb{R})_+ \ \forall\, A \in M_n\big([-1,1]\big)_+,\ n \geqslant 1\}$$

is a Choquet simplex with closed set of extreme points

$$\{x^n : n \geqslant 0\} \cup \{1_{\{1\}} - 1_{\{-1\}}, 1_{\{-1,1\}}\}.$$

□

Horn's theorem

Theorem 4 (Horn, 1969)

If $f : (0, \infty) \to \mathbb{R}$ *is continuous and such that* $f[A] \in M_n(\mathbb{R})_+$ *for all* $A \in M_n\big((0, \infty)\big)_+$, *where* $n \geqslant 3$, *then*

(i) $f \in C^{n-3}\big((0, \infty)\big)$ *and*

(ii) $f^{(k)}(x) \geqslant 0$ *for all* $k = 0, \ldots, n-3$ *and all* $x > 0$.

If, further, $f \in C^{n-1}\big((0, \infty)\big)$, *then* $f^{(k)}(x) \geqslant 0$ *for all* $k = 0, \ldots, n-1$ *and all* $x > 0$.

Lemma 5

If $f : D(0, \rho) \to \mathbb{R}$ *is analytic, where* $\rho > 0$, *and such that* $f[A] \in M_n(\mathbb{R})_+$ *whenever* $A \in M_n\big((0, \rho)\big)_+$ *has rank at most one, then the first* n *non-zero Taylor coefficients of* f *are strictly positive.*

A theorem for fixed dimension

Theorem 6 (B–G–K–P, 2015)

Let $\rho > 0$, $n \geqslant 1$ *and*

$$f(z) = \sum_{j=0}^{n-1} c_j z^j + c' z^m,$$

where $\mathbf{c} = (c_0, \ldots, c_{n-1}) \in \mathbb{R}^n$, $c' \in \mathbb{R}$ *and* $m \geqslant 0$. *The following are equivalent.*

(i) $f[A] \in M_n(\mathbb{C})_+$ *for all* $A \in M_n\big(\overline{D}(0, \rho)\big)_+$.

(ii) *Either* $\mathbf{c} \in \mathbb{R}_+^n$ *and* $c' \in \mathbb{R}_+$, *or* $\mathbf{c} \in (0, \infty)^n$ *and* $c' \geqslant -\mathfrak{C}(\mathbf{c}; m, \rho)^{-1}$, *where*

$$\mathfrak{C}(\mathbf{c}; m, \rho) := \sum_{j=0}^{n-1} \binom{m}{j}^2 \binom{m-j-1}{n-j-1}^2 \frac{\rho^{m-j}}{c_j}.$$

(iii) $f[A] \in M_n(\mathbb{R})_+$ *for all* $A \in M_n\big((0, \rho)\big)_+$ *with rank at most one.*

A determinant identity with Schur polynomials

For all $t \in \mathbb{R}$, $n \geqslant 1$ and $\mathbf{c} = (c_0, \ldots, c_{n-1}) \in \mathbb{F}^n$, let the polynomial

$$p(z; t, \mathbf{c}, m) := t(c_0 + \cdots + c_{n-1} z^{n-1}) - z^m \qquad (z \in \mathbb{C}).$$

Given a non-increasing n-tuple $\mathbf{k} = (k_n \geqslant \cdots \geqslant k_1) \in \mathbb{Z}_+^n$, let

$$\sigma_{\mathbf{k}}(x_1, \ldots, x_n) := \frac{\det(x_i^{k_j + j - 1})}{\det(x_i^{j-1})} \quad \text{and} \quad V_n(x_1, \ldots, x_n) := \det(x_i^{j-1}).$$

Theorem 7

Let $m \geqslant n \geqslant 1$. If $\mu(m, n, j) := (m - n + 1, 1, \ldots, 1, 0, \ldots, 0)$, where there are $n - j - 1$ ones and j zeros, and $\mathbf{c} \in (\mathbb{F}^\times)^n$ then, for all $\mathbf{u}$, $\mathbf{v} \in \mathbb{F}^n$,

$$\det p[\mathbf{u}\mathbf{v}^T; t, \mathbf{c}, m] = t^{n-1} V_n(\mathbf{u}) V_n(\mathbf{v}) \Big(t - \sum_{j=0}^{n-1} \frac{\sigma_{\mu(m,n,j)}(\mathbf{u}) \sigma_{\mu(m,n,j)}(\mathbf{v})}{c_j} \Big) \prod_{j=0}^{n-1} c_j.$$

Proof that (ii) $\implies$ (i) in Theorem 6 when $m \geqslant n$

By the Schur product theorem, it suffices to assume that $-\mathfrak{C}(\mathbf{c}; m, \rho)^{-1} \leqslant c' < 0 < c_0, \ldots, c_{n-1}$.

The key step is to show that $f[A] \in M_n(\mathbb{C})_+$ if $A \in M_n\big(\overline{D}(0,\rho)\big)_+$ has rank at most one.

Given this, the proof concludes by induction. The $n = 1$ case is immediate, since every 1×1 matrix has rank at most one.

Assume the $n - 1$ case holds. Since

$$f(z) = \sum_{j=0}^{n-1} c_j z^j + c' z_m = |c'|\big(|c'|^{-1}(c_0 + \cdots + c_{n-1} z^{n-1}) - z^m\big) = |c'|\, p(z; |c'|^{-1}, \mathbf{c}, m),$$

it suffices to show $p[A] := p[A; |c'|^{-1}, \mathbf{c}, m] \in M_n(\mathbb{C})_+$ if $A \in M_n(\mathbb{C})_+$.

Two lemmas

Lemma 8 (FitzGerald–Horn, 1977)

Given $A = (a_{ij}) \in M_n(\mathbb{C})_+$, set $\mathbf{z} \in \mathbb{C}^n$ to equal $(a_{in}/\sqrt{a_{nn}})$ if $a_{nn} \neq 0$ and to be the zero vector otherwise. Then $A - \mathbf{z}\mathbf{z}^* \in M_n(\mathbb{C})_+$ and the final row and column of this matrix are zero.

If $A \in M_n\big(\overline{D}(0,\rho)\big)_+$ then $\mathbf{z}\mathbf{z}^* \in M_n\big(\overline{D}(0,\rho)\big)_+$, since

$$\begin{vmatrix} a_{ii} & a_{in} \\ \overline{a_{in}} & a_{nn} \end{vmatrix} = a_{ii}a_{nn} - |a_{in}|^2 \geqslant 0.$$

Lemma 9

If $F : \mathbb{C} \to \mathbb{C}$ is differentiable then

$$F(z) = F(w) + \int_0^1 (z-w)F'(\lambda z + (1-\lambda)w)\,\mathrm{d}\lambda \qquad (z, w \in \mathbb{C}).$$

Proof that (ii) $\implies$ (i) continued

By Lemmas 8 and 9,

$$p[A] = p[\mathbf{z}\mathbf{z}^*] + \int_0^1 (A - \mathbf{z}\mathbf{z}^*) \circ p'[\lambda A + (1-\lambda)\mathbf{z}\mathbf{z}^*] \, \mathrm{d}\lambda.$$

Now,

$$p'(z; |c'|^{-1}, \mathbf{c}, m) = m \, p(z; |c'|^{-1}/m, (c_1, \ldots, (n-1)c_{n-1}), m-1)$$

so the integrand is positive semidefinite, by the inductive assumption, since

$$m \, \mathfrak{C}\big((c_1, \ldots, (n-1)c_{n-1}); m-1, \rho\big) \leqslant \mathfrak{C}(\mathbf{c}; m, \rho).$$

Hence $p[A]$ is positive semidefinite, as required.

Hankel matrices

Hankel matrices

Let μ be a measure on $\mathbb{R}$ with moments of all orders, and let

$$s_n = s_n(\mu) := \int_R x^n \mu(\mathrm{d}x) \qquad (n \geqslant 0).$$

The *Hankel matrix associated with* μ is

$$H_\mu := \begin{pmatrix} s_0 & s_1 & s_2 & \dots \\ s_1 & s_2 & s_3 & \dots \\ s_2 & s_3 & s_4 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = (s_{i+j})_{i,j \geqslant 0}.$$

The Hamburger moment problem

Theorem 10 (Hamburger)

A sequence $(s_n)_{n \geqslant 0}$ is the moment sequence for a positive Borel measure on $\mathbb{R}$ if and only if the associated Hankel matrix is positive semidefinite.

Corollary 11

A map f preserves positivity when applied entrywise to Hankel matrices if and only if it maps moment sequences to themselves: given any positive Borel measure μ,

$$f\big(s_n(\mu)\big) = s_n(\nu) \qquad (n \geqslant 0)$$

for some positive Borel measure ν.

Preserving positivity for Hankel matrices

Theorem 12 (B–G–K–P, 2016)

Let $f : \mathbb{R} \to \mathbb{R}$. The following are equivalent.

1. *The function f maps moment sequences of measures supported on $[-1, 1]$ into themselves.*
2. *If A is an $N \times N$ positive semidefinite Hankel matrix then $f[A]$ is positive semidefinite, for all $N \geqslant 1$.*
3. *If A is an $N \times N$ positive semidefinite matrix then $f[A]$ is positive semidefinite, for all $N \geqslant 1$.*
4. *The function f is the restriction to $\mathbb{R}$ of an entire function $F(z) = \sum_{n=0}^{\infty} a_n z^n$, with $a_n \geqslant 0$ for all $n \geqslant 0$.*

Total non-negativity and total positivity

A matrix is *totally non-negative* if each of its minors is non-negative.
A matrix is *totally positive* if each of its minors is positive.
These were first investigated by Fekete (1910s) then by Gantmacher and Krein and by Schoenberg (1930s). There are of much recent interest.

Example

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{pmatrix}$$

Remark

A totally non-negative matrix which is symmetric is positive semidefinite, by Sylvester's criterion.

Preserving positivity for symmetric TN matrices

Theorem 13 (B–G–K–P, 2017)

Let $F : [0, \infty) \to \mathbb{R}$ and let N be a positive integer. The following are equivalent.

1. *F preserves total non-negativity entrywise on symmetric $N \times N$ matrices.*
2. *F is either a non-negative constant or*
   - (a) $(N = 1)$ $F(x) \geqslant 0$;
   - (b) $(N = 2)$ *F is non-negative, non-decreasing, and multiplicatively mid-convex, i.e.,*

$$F(\sqrt{xy})^2 \leq F(x)F(y) \qquad \text{for all } x, y \in [0, \infty);$$

   - (c) $(N = 3)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha \geqslant 1$;
   - (d) $(N = 4)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha \in \{1\} \cup [2, \infty)$;
   - (e) $(N \geqslant 5)$ $F(x) = cx$ *for some* $c > 0$.

Preserving positivity for symmetric TP matrices

Theorem 14 (B–G–K–P, 2017)

Let $F : (0, \infty) \to \mathbb{R}$ and let N be a positive integer. The following are equivalent.

1. *F preserves total positivity entrywise on symmetric $N \times N$ matrices.*
2. *The function F satisfies*
   - (a) $(N = 1)$ $F(x) > 0$;
   - (b) $(N = 2)$ *F is positive, increasing, and multiplicatively mid-convex, i.e.,*

$$F(\sqrt{xy})^2 \leq F(x)F(y) \qquad \text{for all } x, y \in (0, \infty);$$

   - (c) $(N = 3)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha \geqslant 1$;
   - (d) $(N = 4)$ $F(x) = cx^\alpha$ *for some* $c > 0$ *and some* $\alpha \in \{1\} \cup [2, \infty)$.
   - (e) $(N \geqslant 5)$ $F(x) = cx$ *for some* $c > 0$.

The end!

Thank you for your attention

Dziękuję za uwagę

References

A. Belton, D. Guillot, A. Khare and M. Putinar, *Matrix positivity preservers in fixed dimension. I*, Adv. Math. **298** (2016) 325–368.

A. Belton, D. Guillot, A. Khare and M. Putinar, *Matrix positivity preservers in fixed dimension*, C. R. Math. Acad. Sci. Paris **354** (2016) 143–148.

A. Belton, D. Guillot, A. Khare and M. Putinar, *Moment-sequence transforms*, arXiv:1610.05740.

A. Belton, D. Guillot, A. Khare and M. Putinar, *Total-positivity preservers*, arXiv:1711.10468

J. P. R. Christensen and P. Ressel, *Functions operating on positive definite matrices and a theorem of Schoenberg*, Trans. Amer. Math. Soc. **243** (1978) 89–95.

References

C. H. FitzGerald and R. A. Horn, *On fractional Hadamard powers of positive definite matrices*, J. Math. Anal. Appl. **61** (1977), 633–642.

R. A. Horn, *The theory of infinitely divisible matrices and kernels*, Trans. Amer. Math. Soc. **136** (1969) 269–286.

I. J. Schoenberg, *Positive definite functions on spheres*, Duke Math. J. **9** (1942) 96–108.

J. Schur, *Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen*, J. Reine Angew. Math. **140** (1911) 1–28.