Preservers for positive-semidefinite and totally non-negative matrices

Alexander Belton

Department of Mathematics and Statistics
Lancaster University, United Kingdom
a.belton@lancaster.ac.uk

Portsea Maths Research Webinar
22nd July 2020
Structure of the talk

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

1. Theorems of Schur, Schoenberg and Horn
2. Positivity preservation
   1. For fixed dimension
   2. For moment matrices
   3. For totally non-negative symmetric matrices
   4. For totally non-negative matrices
3. Hankel and Toeplitz kernels
   1. Positivity preservers for continuous Hankel kernels
   2. Positivity preservers for Pólya frequency functions
Positive semidefiniteness

**Definition**

Let $A \in M_d(\mathbb{C})$. The following are equivalent.

- $z^* A z \geq 0$ for all $z \in \mathbb{C}^d$.
- $A = A^*$ and $\sigma(A) \subseteq \mathbb{R}_+$.
- $A = U \text{diag}(\lambda_1, \ldots, \lambda_d) U^*$, with $\lambda \in \mathbb{R}_+^d$ and $U^* U = U U^* = I$.
- $A = \sum_{i \in [d]} \lambda_i e_i e_i^*$, where $\lambda_i \geq 0$ and $e_i^* e_j = 1_{i=j}$ if $i, j \in [d]$.
- $A = B^* B$, where $B \in M_n(\mathbb{C})$.

If these conditions hold, then $A$ is **positive semidefinite** and we write $A \succeq 0$.

**Remark**

The collection of $d \times d$ positive semidefinite matrices is a closed cone: stable under addition, linear homotheties and pointwise limits.
The Schur product theorem

**Theorem 1 (Schur, 1911)**

If \( A, B \succeq 0 \) then \( A \circ B \succeq 0 \), where the Schur product is such that

\[
(A \circ B)_{ij} := a_{ij} b_{ij} \quad \text{for all } i, j.
\]

**Proof 1.**

Note that \( A^T \) is positive semidefinite whenever \( A \) is, and so

\[
z^*(A \circ B)z = \text{tr} \left( \text{diag}(z)^* B \text{diag}(z) A^T \right)
= \text{tr} \left( (A^T)^{1/2} \text{diag}(z)^* B \text{diag}(z) (A^T)^{1/2} \right).
\]

**Proof 2.**

Note that

\[
(ee^*) \circ (ff^*) = (e \circ f)(e \circ f)^*,
\]

so the rank-one case hold, and use linearity.
A question of Pólya and Szegö

**Corollary 2**

If \( f(z) = \sum_{n=0}^{\infty} c_n z^n \), with \( c_n \geq 0 \) for all \( n \), then

\[
 f[A] := \sum_{n=0}^{\infty} c_n A^n = \begin{pmatrix} f(a_{11}) & \cdots & f(a_{1d}) \\ \vdots & \ddots & \vdots \\ f(a_{d1}) & \cdots & f(a_{dd}) \end{pmatrix} \geq 0
\]

for any \( A \geq 0 \) such that each \( a_{ij} \) lies in the disc of convergence for \( f \).

**Observation**

A function \( f \) with such a power-series expansion is said to be *absolutely monotonic*. Absolutely monotonic functions preserve positivity independent of the dimension \( d \).

**Question (Pólya–Szegö, 1925)**

Are there any other functions with this property?
Schoenberg’s theorem

**Theorem 3 (Schoenberg, 1942)**

If \( f : [-1, 1] \rightarrow \mathbb{R} \) is

(i) continuous and

(ii) such that \( f[A] \geq 0 \) for every \( A \geq 0 \) with entries in \([-1, 1]\) and of any size,

then \( f \) is absolutely monotonic:

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

**Theorem 4 (Rudin, 1959)**

*The hypothesis of continuity is not required in Schoenberg’s result.*
Horn’s theorem

Theorem 5 (Horn, 1969)

If \( f : (0, \infty) \rightarrow \mathbb{R} \) is continuous and such that \( f[A] \geq 0 \) for every \( d \times d \) positive-semidefinite matrix \( A \), where \( d \geq 3 \), then

(i) \( f \) is \( d - 3 \)-times continuously differentiable and

(ii) \( f^{(k)}(x) \geq 0 \) for all \( k = 0, \ldots, d - 3 \) and all \( x > 0 \).

If, further, \( f \) is \( d - 1 \)-times differentiable, then \( f^{(k)}(x) \geq 0 \) for all \( k = 0, \ldots, d - 1 \) and all \( x > 0 \).

Observation (Guillot–Khare–Rajaratnam)

To obtain Horn’s result, it suffices to consider only matrices having the form \( A = a1 + uu^T \), where \( a > 0 \) and \( u_i \geq 0 \) whenever \( i \in [d] \).
Proposition 6 (B–G–K–P, 2016)

If \( f : D(0, \rho) \to \mathbb{R} \) is analytic, where \( \rho > 0 \), and such that \( f[A] \geq 0 \) whenever \( A \) is a \( d \times d \) positive-semidefinite matrix of rank one with entries in \((0, \rho)\), then the first \( d \) non-zero Maclaurin coefficients of \( f \) are strictly positive.

Proof.

Suppose the first \( d \) non-zero Maclaurin coefficients are \( c_{n_1}, \ldots, c_{n_d} \). Let \( u^T = (u_1, \ldots, u_d) \) have distinct entries in \((0, \sqrt{\rho})\), and note that \((u_i^{n_j})\) is a generalised Vandermonde matrix, so invertible. Hence \( \{u^{n_1}, \ldots, u^{n_d}\} \) is linearly independent; let \( \{v_1, \ldots, v_d\} \subseteq \mathbb{R}^d \) be such that \( v_i^T u^{n_j} = \mathbb{1}_{i=j} \).

Then, for any \( i \in [d] \),

\[
0 \leq \varepsilon^{-n_i} v_i^T f[\varepsilon uu^T] v_i = c_{n_i} + \sum_{j > n_i} c_j \varepsilon^{j-n_i} (v_i^T u^{n_j})^2 \rightarrow c_{n_i} \text{ as } \varepsilon \rightarrow 0^+.
\]

\[\square\]
The following are equivalent.

(i) \( f[A] \succeq 0 \) for every \( d \times d \) positive-semidefinite matrix \( A \) with entries in \( D(0, \rho) \).

(ii) Either \( c \in \mathbb{R}_+^d \) and \( c' \in \mathbb{R}_+ \), or \( c \in (0, \infty)^d \) and \( c' \succeq -C(c; m, \rho)^{-1} \), where

\[
C(c; m, \rho) := \sum_{j=0}^{d-1} \binom{m}{j}^2 \binom{m-j-1}{d-j-1}^2 \rho^{m-j} \frac{c_j}{c_j}.
\]

(iii) \( f[A] \succeq 0 \) for every \( d \times d \) positive-semidefinite matrix with entries in \( (0, \rho) \) and rank at most one.
Hankel matrices

Moment matrices

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

$$s_n = s_n(\mu) := \int_{\mathbb{R}} x^n \, \mu(dx) \quad (n \geq 0).$$

The **Hankel matrix associated with** $\mu$ is

$$H_\mu := \begin{pmatrix} s_0 & s_1 & s_2 & \cdots \\ s_1 & s_2 & s_3 & \cdots \\ s_2 & s_3 & s_4 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = (s_{i+j})_{i,j \geq 0}.$$
The Hamburger moment problem

Theorem 8 (Hamburger)

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

Corollary 9

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 \( \mu \),

\[
f(s_n(\mu)) = s_n(\nu) \quad (n \geq 0)
\]

for some positive Borel measure \( \nu \).
Theorem 10 (B–G–K–P, 2016)

Let \( f : \mathbb{R} \rightarrow \mathbb{R} \). The following are equivalent.

1. The function \( f \) maps the set of moment sequences of measures supported on \([-1, 1]\) into itself.
2. \( f[A] \succeq 0 \) whenever \( A \) is a positive-semidefinite Hankel matrix of any size.
3. \( f[A] \succeq 0 \) whenever \( A \) is a positive-semidefinite matrix of any size.
4. The function \( f \) is absolutely monotonic:

\[
    f(x) = \sum_{n=0}^{\infty} c_n x^n \quad \text{for all } x \in \mathbb{R},
\]

with \( c_n \geq 0 \) for all \( n \).
**Total non-negativity**

**Definition**
A matrix is *totally non-negative* if each of its minors is non-negative.

**Remark**
This concept was first investigated by Fekete (1910s) then by Gantmacher and Krein and by Schoenberg (1930s).

**Example**
\[
\begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 \\
1 & 2 & 4 & 8 & 16 \\
1 & 3 & 9 & 27 & 81 \\
1 & 4 & 16 & 64 & 256 \\
\end{pmatrix}
\]

Let \( f : \mathbb{R}_+ \rightarrow \mathbb{R} \) and let \( d \geq 1 \). The following are equivalent.

1. \( f[A] \) is totally non-negative whenever \( A \) is a totally non-negative symmetric \( d \times d \) matrix.

2. \( f \) is either a non-negative constant or
   (a) \( (d = 1) \) \( f(x) \geq 0 \);
   (b) \( (d = 2) \) \( f \) is non-negative, non-decreasing, and multiplicatively mid-convex, that is,

\[
    f(\sqrt{xy})^2 \leq f(x)f(y) \quad \text{for all } x, y \in [0, \infty),
\]

which implies that \( f \) is continuous on \((0, \infty)\);
   (c) \( (d = 3) \) \( f(x) = cx^{\alpha} \) for some \( c > 0 \) and some \( \alpha \geq 1 \);
   (d) \( (d = 4) \) \( f(x) = cx^{\alpha} \) for some \( c > 0 \) and some \( \alpha \in \{1\} \cup [2, \infty) \);
   (e) \( (d \geq 5) \) \( f(x) = cx \) for some \( c > 0 \).
Theorem 12 (B–G–K–P, 2020)

Let $f : (0, \infty) \to \mathbb{R}$ and let $d \geq 1$. The following are equivalent.

1. $f[A]$ is totally non-negative whenever $A$ is a totally non-negative $d \times d$ matrix.

2. $f$ is either a non-negative constant or
   (a) $(d = 1)$ $f(x) \geq 0$;
   (b) $(d = 2)$ $f(x) = cx^\alpha$ for some $c > 0$ and some $\alpha \geq 0$;
   (c) $(d = 3)$ $f(x) = cx^\alpha$ for some $c > 0$ and some $\alpha \geq 1$.
   (d) $(d \geq 4)$ $f(x) = cx$ for some $c > 0$.

Sketch proof for $d = 2$

Looking at $\det f[A]$ and $\det f[B]$, where

$$A = \begin{pmatrix} x & \sqrt{xy} \\ \sqrt{xy} & y \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} x & y \\ y & x \end{pmatrix} \quad (x \geq y \geq 0)$$

gives that $f$ is multiplicatively mid-point convex and non-decreasing.
Sketch proof for $d = 2$ (continued).

Furthermore, looking at $\det f[C]$ and $\det f[D]$, where

$$C = \begin{pmatrix} x & xy \\ 1 & y \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} xy & x \\ y & 1 \end{pmatrix}$$

to see that

$$f(xy)f(1) = f(x)f(y) \quad \text{for all} \quad x, y \geq 0.$$

Normalising shows that $f$ satisfies the multiplicative functional equation, and thus $f(x) = f(1)x^\alpha$ for some $\alpha$. As $f$ is non-decreasing, so $\alpha \geq 0$. \hfill \square

Hint for $d = 3$

Look at $\det f[E]$, where

$$E = \begin{pmatrix} \sqrt{2} & 1 & 0 \\ 1 & \sqrt{2} & 1 \\ 0 & 1 & \sqrt{2} \end{pmatrix}.$$
Preserving positivity for TN Hankel matrices

**Theorem 13 (B–G–K–P, 2020)**

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

- $f[A]$ is totally non-negative for every totally non-negative Hankel matrix $A$ of any size.
- $f[A] \geq 0$ for every totally non-negative Hankel matrix $A$ of any size.
- $f(x) = \sum_{n=0}^{\infty} c_n x^n$ for all $x \in (0, \infty)$, where $c_n \geq 0$ for all $n$, and $0 \leq f(0) \leq c_0$.

**Theorem 14 (B–G–K–P, 2020)**

If $f : \mathbb{R}_+ \to \mathbb{R}$ preserves total non-negativity for $d \times d$ Hankel matrices, then $f$ is $d - 3$-times continuously differentiable, with $f$, $f'$, $\ldots$, $f^{(d-3)}$ non-negative, on $(0, \infty)$. If $f$ is analytic then the first $d$ non-zero Maclaurin coefficients for $f$ are positive.
Hankel and Toeplitz kernels

Definition 15

If $X$ and $Y$ are totally ordered sets, then the kernel $K : X \times Y \to \mathbb{R}$ is \textit{positive semidefinite} or \textit{totally non-negative} if and only if the $d \times d$ matrix $(K(x_i, y_j))$ has the same property for every choice of $x_1 < \cdots < x_d$ and $y_1 < \cdots < y_d$, and all $d \geq 1$.

Definition 16

If $X \subseteq \mathbb{R}$ then $K : X \times X \to \mathbb{R}$ is Hankel if there exists $f : \mathbb{R} \to \mathbb{R}$ such that

$$K(x, y) = f(x + y) \quad \text{for all } x, y \in X,$$

or Toeplitz if there exists $f : \mathbb{R} \to \mathbb{R}$ such that

$$K(x, y) = T_f(x, y) := f(x - y) \quad \text{for all } x, y \in X.$$
Proposition 17 (Bernstein, Widder)

Let $X \subseteq \mathbb{R}$ be an open interval and let $K : X \times X \rightarrow \mathbb{R}$ be a continuous Hankel kernel. The following are equivalent.

- $K$ is totally non-negative.
- $K$ is positive semidefinite.
- $K : (x, y) \mapsto \int_{\mathbb{R}} e^{(x+y)u} \, d\sigma(u)$ for some non-decreasing function $\sigma$.

Theorem 18 (B–G–K–P, 2020)

Let $X$ be an interval containing more than one point and let $f : \mathbb{R}^+ \rightarrow \mathbb{R}$. The following are equivalent.

- $f \circ K$ is totally non-negative for every totally non-negative continuous Hankel kernel $K : X \times X \rightarrow \mathbb{R}$.
- $f \circ K$ is positive semidefinite for every TN cts Hankel kernel $K$ on $X$.
- $f(x) = \sum_{n=0}^{\infty} c_n x^n$, with $c_n \geq 0$ for all $n$, on $(0, \infty)$, and $f(0) \geq 0$. 
### Definition 19

The map $\Lambda : \mathbb{R} \to \mathbb{R}$ is a **Pólya frequency function** if it is Lebesgue integrable, non-zero at two or more points and such that the Toeplitz kernel

$T_\Lambda : \mathbb{R} \times \mathbb{R} \to \mathbb{R}; \ (x, y) \mapsto \Lambda(x - y)$

is totally non-negative.

### Example (Gaussian kernels)

For any $\kappa > 0$,

$G_\kappa : \mathbb{R} \times \mathbb{R}; \ (x, y) \mapsto \exp(-\kappa(x - y)^2)$

is the totally non-negative kernel corresponding to the Pólya frequency function $x \mapsto \exp(-\kappa x^2)$. 
Theorem 20 (B–G–K–P, 2020)

Let $f : (0, \infty) \to \mathbb{R}$. Each of the following statements implies the next.

- $f(x) = cx^\alpha$ for some $c > 0$ and $\alpha > 0$.
- $f \circ dG_\kappa$ is totally non-negative for all $d > 0$ and $\kappa > 0$.
- $f \circ dG_1$ is totally non-negative for all $d > 0$.
- $f$ is non-negative, non-decreasing and continuous.

Remark

 Quite wild preservers of total non-negativity exist for Gaussian kernels: since $x \mapsto \exp(-\beta|x|)$ is a Pólya frequency function for any $\beta > 0$, the map

$$f : (0, \infty) \to \mathbb{R}; \ x \mapsto \exp(-\sqrt{-\log \max\{x, 1\}})$$

is such that $f \circ G_\kappa$ is totally non-negative for all $\kappa > 0$. 
Theorem 21 (B–G–K–P, 2020)
There exists a Pólya frequency function $M$ such that $M^n$ is not a Pólya frequency function for any integer $n \geq 2$.

Proof.
Let $M(x) = 2e^{-|x|} - e^{-2|x|}$ for all $x$, then examine the reciprocals of the bilateral Laplace transforms of $M^n$:
- for $n = 1$, this is a polynomial with real roots, so belongs to the Laguerre–Pólya class (entire with real zeros);
- for $n \geq 2$, this is a rational function with simple poles, so does not belong to the Laguerre–Pólya class.

Corollary 22 (B–G–K–P, 2020)
Let $f : \mathbb{R}_+ \to \mathbb{R}$. Then $f \circ \Lambda$ is a Pólya frequency function for every Pólya frequency function $\Lambda$ if and only if $f(x) = cx$ for some $c > 0$. 
Thank you for your attention
References


References


