A quantitative form of Schoenberg’s theorem in fixed dimension

Alexander Belton
(Joint work with Dominique Guillot, Apoorva Khare and Mihai Putinar)

Department of Mathematics and Statistics
Lancaster University, United Kingdom
a.belton@lancaster.ac.uk
http://www.maths.lancs.ac.uk/~belton/

Iowa State University
25th July 2017
Structure of the talk

- The Hadamard product
- Positive definiteness
- Schur’s theorem
- Schoenberg’s theorem
- Horn’s theorem
- A new theorem and its proof
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

$$(AB)_{ij} := \sum_{k=1}^{n} a_{ik} b_{kj} \quad \text{(standard)}$$

and

$$(A \circ B)_{ij} := a_{ij} b_{ij} \quad \text{(Hadamard)}.$$
Positive definiteness. I

**Definition**

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

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

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

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

Symmetry

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

$$0 \leq (x^*Ax)^T = y^*A^Ty$$

for all $y = \bar{x} \in \mathbb{C}^n$.

Hermitianity

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

$$x^*A^*x = (x^*Ax)^* = x^*Ax \quad (x \in \mathbb{C}^n).$$
Non-negative eigenvalues

If

\[ A \in M_n(\mathbb{C})_+ = U^* \text{diag}(\lambda_1, \ldots, \lambda_n)U, \]

where \( U \in M_n(\mathbb{C}) \) is unitary, then \( \lambda_i = (U^* e_i)^* A (U^* e_i) \geq 0 \).

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

\[ B = A^{1/2} = U^* \text{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})U. \]

Corollary

If \( A \in M_n(\mathbb{C})_+ \) then

\[ \det A = \lambda_1 \cdots \lambda_n \geq 0 \]

and

\[ \text{tr } A = \lambda_1 + \cdots + \lambda_n \geq 0. \]
A characterisation of positive semidefiniteness

Definition

A principal minor of an $n \times n$ matrix is the determinant of the $k \times k$ matrix formed by deleting $n - k$ rows and the corresponding $n - k$ columns.

Theorem 1 (Sylvester’s criterion)

A Hermitian matrix $A \in M_n(\mathbb{C})$ is positive semidefinite if and only if each of its principal minors is non-negative.

Proof.

For the converse, proceed by induction. If the $n \times n$ matrix $A$ is not positive semidefinite but all of its principal minors are non-negative then it has two negative eigenvalues with unit eigenvectors $\mathbf{x}$ and $\mathbf{y}$. The unit vector $\mathbf{z} = \lambda \mathbf{x} + \mu \mathbf{y}$ can be taken orthogonal to $(0, \ldots, 0, 1) \perp$ and is such that

$$\mathbf{z}^* A \mathbf{z} = |\lambda|^2 \mathbf{x}^* A \mathbf{x} + |\mu|^2 \mathbf{y}^* A \mathbf{y} < 0.$$
The Schur product theorem

**Theorem 2 (Schur, 1911)**

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

**Proof.**

\[
\mathbf{x}^*(A \circ B)\mathbf{x} = \text{tr}(\text{diag}(\mathbf{x})^* B \text{diag}(\mathbf{x})A^T)
\]

\[
= \text{tr}((A^T)^{1/2} \text{diag}(\mathbf{x})^* B \text{diag}(\mathbf{x})(A^T)^{1/2}).
\]
Schoenberg’s theorem

Corollary 3
If \( f(z) = \sum_{n=0}^{\infty} a_n z^n \) is analytic on \( K \) and \( a_n \geq 0 \) for all \( n \geq 0 \) then
\[
f[A] := (f(a_{ij})) \in M_n(\mathbb{C})_+ \quad \text{for all } A = (a_{ij}) \in M_n(K)_+ \text{ and all } n \geq 1.
\]

Theorem 4 (Schoenberg, 1942)
If \( f : [-1, 1] \rightarrow \mathbb{R} \) is
(i) continuous and
(ii) such that \( f[A] \in M_n(\mathbb{R})_+ \) for all \( A \in M_n([-1, 1])_+ \) and all \( n \geq 1 \)
then \( f \) is absolutely monotonic:
\[
f(x) = \sum_{n=0}^{\infty} a_n x^n \quad \text{for all } x \in [-1, 1], \quad \text{where } a_n \geq 0 \text{ for all } n \geq 0.
\]
Horn’s theorem

**Theorem 5 (Horn, 1969)**

If \( f : (0, \infty) \to \mathbb{R} \) is continuous and such that \( f[A] \in M_n(\mathbb{R})_+ \) whenever \( A \in M_n((0, \infty))_+ \) then

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

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

If, further, \( f \in C^{n-1}((0, \infty)) \), then

\[ f^{(k)}(x) \geq 0 \quad \text{for all } k = 0, \ldots, n-1 \text{ and } x > 0. \]

**Remark (Guillot–Khare–Rajaratnam, 2014)**

For Theorem 5, it suffices to suppose \( \rho > 0 \) and \( f : (0, \rho) \to \mathbb{R} \) is such that

\[ f[a1_n + xx^T] \in M_n(\mathbb{R})_+ \quad \text{for all } a \in [0, \rho) \text{ and } x \in [0, (\rho - a)^{1/2}). \]
Theorem 6 (B–G–K–P, 2015)

Let $\rho > 0$, $n \geq 1$, $\mathbf{c} = (c_0, \ldots, c_{n-1}) \in \mathbb{R}^n$ and

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

where $c' \in \mathbb{R}$ and $m \geq 0$. The following are equivalent.

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

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

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

$$\mathcal{C}(\mathbf{c}; m, \rho) := \sum_{j=0}^{n-1} \binom{m}{j}^2 \binom{m-j-1}{n-j-1}^2 \rho^{m-j} c_j.$$
Proof of Theorem 6 with \( m \geq n \)

**Lemma 7 (B–G–K–P, 2015)**

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

**Proof of Theorem 6**

Ir is immediate that (i) implies (ii).

To see (ii) implies (iii), note that Lemma 7 gives that \( c_0, \ldots, c_{n-1} \geq 0 \), so we assume \( c' < 0 < c_0, \ldots, c_{n-1} \) and show that \( c' \geq -\mathfrak{c}(c; m, \rho)^{-1} \). If

\[
p(z; t, c, m) := t(c_0 + \cdots + c_{n-1}z^{n-1}) - z^m \quad (z \in \mathbb{C}, \ t \in \mathbb{R})
\]

then (ii) implies that

\[
\det p[uu^T; |c'|^{-1}, c, m] \geq 0 \quad \text{for all } u \in (0, \sqrt{\rho})^n.
\]
A determinant expressed with Schur polynomials

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

$$s_k(x_1, \ldots, x_n) := \frac{\det(x_i^{k_j + j - 1})}{\det(x_i^{j - 1})}$$

and

$$V_n(x_1, \ldots, x_n) := \det(x_i^{j - 1}).$$

**Theorem 8**

Let $m \geq n \geq 1$. 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 $c \in (\mathbb{F}^\times)^n$ then, for all $u, v \in \mathbb{F}^n$,

$$\det p[uv^T; t, c, m] = t^{n-1} V_n(u) V_n(v) \left( t - \sum_{j=0}^{n-1} \frac{s_{\mu(m, n, j)}(u) s_{\mu(m, n, j)}(v)}{c_j} \right) \prod_{j=0}^{n-1} c_j.$$
Proof of Theorem 6, $m \geq n$, (ii) $\implies$ (iii) (ctd.)

It follows from the previous working that

$$0 \leq \det p[uu^T; |c'|^{-1}, c, m]$$

$$= |c'|^{1-n} V_n(u)^2 \left( |c'|^{-1} - \sum_{j=0}^{n-1} \frac{s_\mu(m,n,j)(u)^2}{c_j} \right) \prod_{j=0}^{n-1} c_j.$$

Taking $u_j = \sqrt{\rho} \delta_j$, where the $\delta_j$ are distinct and tend to 1 from below, it follows that

$$|c'|^{-1} \geq \sum_{j=0}^{n-1} s_\mu(m,n,j)(1, \ldots, 1)^2 \rho^{m-j} = \mathcal{C}(c; m, \rho).$$
Proof of Theorem 6, \( m \geq n \), \((iii) \implies (i)\)

It remains to prove that \((iii)\) implies \((i)\). By Schur’s Theorem 2, we may assume that \(-C(c; m, \rho)^{-1} \leq c' < 0 < c_0, \ldots, c_{n-1}.

**Key step**

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

Given this, the proof concludes by induction. The \(n = 1\) case is immediate, since every \(1 \times 1\) matrix has rank at most one. Assume the \(n - 1\) case holds. Since

\[ f(z) = |c'| p(z; |c'|^{-1}, c, m), \]

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

Lemma 9 (FitzGerald–Horn, 1977)

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

Furthermore, if $A \in M_n(\overline{D}(0, \rho))_+$ then $zz^* \in M_n(\overline{D}(0, \rho))_+$, since

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

Lemma 10

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

$$
F(z) = F(w) + \int_0^1 (z - w)F'(\lambda z + (1 - \lambda)w)\,d\lambda \quad (z, w \in \mathbb{C}).
$$
Proof of Theorem 6: $m \geq n$, $(iii) \implies (i)$

By Lemmas 9 and 10,

$$p[A] = p[zz^*] + \int_0^1 (A - zz^*) \circ p'[\lambda A + (1 - \lambda)zz^*] \, d\lambda.$$  

Now,

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

which is positive semidefinite, by the inductive assumption, since

$$m \mathcal{C}((c_1, \ldots, (n - 1)c_{n-1}); m - 1, \rho) \leq \mathcal{C}(c; m, \rho).$$

Hence $p[A]$ is positive semidefinite, as required.
Fix \( u \in \mathbb{C}^n \) such that \( A = uu^* \in M_n(\overline{D}(0, \rho))_+ \). For \( k = 1, \ldots, n \), let

\[
C_k := \sum_{j=0}^{k-1} s_\mu(m-n+k,k,j)(1, \ldots, 1)^2 \frac{\rho^{m-n+k-j}}{c_{n-k+j}}
\]

and note that

\[
0 < C_1 = \rho^{m-n+1}/c_{n-1} < \cdots < C_n = \mathcal{C}(c; m, \rho).
\]

We claim that, for \( k = 1, \ldots, n \), every principal \( k \times k \) submatrix of

\[
B = C_k(c_{n-k}1_n + \cdots + c_{n-1}A)^{(k-1)} - A^{(m-n+k)}
\]

is positive semidefinite. When \( k = n \), the matrix \( B \) is equal to \( C_n f[A] \) with \( c' = -\mathcal{C}(c; m, \rho)^{-1} \), and we are done.
Again we proceed by induction. If \( k = 1 \) then

\[
B = C_1 c_{n-1} 1_n - A^{o m-n+1} = \rho^{m-n+1} 1_N - A^{o m-n+1}.
\]

The principal \( 1 \times 1 \) submatrices equal \( \rho^{m-n+1} - |u_i|^2(m-n+1) \), and these are non-negative since \( |u_i|^2 \leq \rho \).

Suppose the \( k - 1 \) case holds. For any non-empty \( k \subseteq \{1, \ldots, n\} \), let \( B_k \) denote the principal submatrix of \( B \) with rows and columns labelled by elements of \( k \), and similarly for \( u_k \).

If \( k \) has cardinality \( j < k \) then

\[
B_k \geq C_k (c_{n-j} A_{k}^{o k-j} + \cdots + c_{n-1} A_{k}^{o k-1}) - A_{k}^{o m-n+k} \\
\geq A_{k}^{o k-j} \circ (C_j(c_{n-j} 1_j + \cdots + c_{n-1} A_{k}^{o j-1}) - A_{k}^{o m-n+j}) \geq 0,
\]

by induction.
Finally, if $k$ has cardinality $k$ then

$$B_k = p[u_k u_k^*; C_k, (c_{n-1}, \ldots, c_{n-k}), m - n + k]$$

and Theorem 8 gives that

$$\det B_k = C_{n-1}^k |V_k(u_k)|^2 \left( C_k - \sum_{j=0}^{k-1} \frac{|s_{\mu(m-n+k,k,j)}(u)|^2}{c_{n-k+j}} \right) \prod_{j=1}^{k} c_{n-j}.$$ 

It now follows from the triangle inequality and the fact that Schur polynomials have non-negative coefficients that $\det B_k \geq 0$.

The end!

Thank you for your attention.


