A new inertia theorem for Stein equations, inertia of invertible hermitian block Toeplitz matrices and matrix orthogonal polynomials

André Ran
VU Amsterdam

Joint work with Leonid Lerer
Technion, Haifa, Israel

Work supported by NWO

Introduction

Invertible hermitian Toeplitz matrix

$$T_n = \left(t_{i-j}\right)_{i,j=0}^{n} = \begin{pmatrix} t_0 & t_{-1} & \cdots & t_{-n} \\ t_1 & t_0 & t_{-1} & \cdot \\ \vdots & \cdot & \cdot & \vdots \\ t_n & \cdots & \cdots & t_0 \end{pmatrix}$$

with $t_i = \bar{t}_{-i}$.

Let

$$T_n \begin{pmatrix} x_{00}^{(n)} \\ x_{10}^{(n)} \\ \vdots \\ x_{n0}^{(n)} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

The n'th orthogonal polynomial is then

$$p_n(\lambda) = \sum_{j=0}^{n} \lambda^{n-j} x_{j0}^{(n)}$$

Krein's Theorem

Given real valued $h \in L_1(T)$, define

$$\langle p, q \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} p(e^{i\theta}) h(e^{i\theta}) \overline{q(e^{i\theta})} \, d\theta.$$

Orthogonalize $1, z, z^2 \cdots$. Then we obtain orthogonal polynomials $p_1(z), p_2(z), \cdots, p_n(z), \cdots$.

Need invertibility of $T_k = \left(t_{i-j}\right)_{i,j=0}^{k}$ $(k = 0, 1, \cdots)$ where t_j are the Fourier coefficients of h.

Theorem 1. (Kreĭn) *Assume* $x_{00}^{(n)} > 0$*. Then*

- $p_n(z)$ *has no zeroes on the unit circle,*

- *the number of positive eigenvalues of* T_{n-1} *equals the number of zeroes of* $p_n(z)$ *inside the unit disc.*

If $x_{00}^{(n)} < 0$*, replace "positive" by "negative".*

Remark: if $h > 0$: Szegö polynomials.

Matrix Generalizations

Consider

$$T_n = \left(t_{i-j}\right)_{i,j=0}^{n}$$

with $t_i = t_{-i}^*$ being $r \times r$ matrices.

Again let

$$T_n \begin{pmatrix} X_{00}^{(n)} \\ X_{10}^{(n)} \\ \vdots \\ X_{n0}^{(n)} \end{pmatrix} = \begin{pmatrix} I_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad (1)$$

and

$$P_n(\lambda) = \sum_{j=0}^{n} \lambda^{n-j} X_{j0}^{(n)}.$$

We assume throughout that T_n and T_{n-1} are invertible.

Matrix Generalization

Theorem 2. *Suppose T_n and T_{n-1} are invertible.*

- *If $X_{00}^{(n)} > 0$ then*

$$\#\{\text{negative eigenvalues of } T_n\} = \#\{\text{zeroes of } P_n \text{ outside the unit disc }\},$$

$$\#\{\text{positive eigenvalues of } T_n\} = r + \#\{\text{zeroes of } P_n \text{ inside the unit disc }\}.$$

- *If $X_{00}^{(n)} < 0$ then*

$$\#\{\text{negative eigenvalues of } T_n\} = r + \#\{\text{zeroes of } P_n \text{ inside the unit disc }\},$$

$$\#\{\text{positive eigenvalues of } T_n\} = \#\{\text{zeroes of } P_n \text{ outside the unit disc }\}.$$

See OT 34 for three different proofs.

For non-definite $X_{00}^{(n)}$ there is no simple relation between $\mathrm{in}\,(T_n)$ and the number of zeroes of $P_n(\lambda)$ inside or outside the unit disc. Examples show that almost every combination is possible, including zeroes of $P_n(\lambda)$ on the unit circle.

However, $\mathrm{in}\,(T_n)$ can be expressed via the first and last column of T_n^{-1}.

Denote the two orthogonal matrix polynomials connected to the first and last block column of T_n^{-1} by

$$
\begin{aligned}
X_0(\lambda) &= \sum_{j=0}^{n} \lambda^j X_{n-j\,0}^{(n)}, \\
X_n(\lambda) &= \sum_{j=0}^{n} \lambda^j X_{n-j\,n}^{(n)}.
\end{aligned}
$$

Main Theorem

Theorem 3. *Suppose* $X_{0n}^{(n)}$ *is invertible, and let* Y *be any matrix for which*

$$Y^*(X_{0n}^{(n)})^{-*}X_{nn}^{(n)}(X_{0n}^{(n)})^{-1}Y < (X_{00}^{(n)})^{-1}.$$

Let $X_0(\lambda)$ *and* $X_n(\lambda)$ *be the two orthogonal matrix polynomials just introduced. Define* $P(\lambda)$ *by*

$$P(\lambda) = \lambda X_0(\lambda)(X_{00}^{(n)})^{-1} + X_n(\lambda)(X_{0n}^{(n)})^{-1}Y,$$

then the number of positive, respectively, negative eigenvalues of T_n *is equal to the number of zeroes of* $P(\lambda)$ *inside, respectively, outside the unit disc (multiplicities taken into account).*

Example

$$T_1 = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}$$

has two positive and two negative eigenvalues. $X_0(\lambda) = \frac{1}{2}\begin{pmatrix} \lambda & 1 \\ 1 & -\lambda \end{pmatrix}$ has zeroes on the unit circle.

It turns out that we can take $Y = \begin{pmatrix} 0 & \varepsilon \\ 0 & 0 \end{pmatrix}$ with $|\varepsilon| > 1$.

Then

$$P(\lambda) = \frac{1}{2}\begin{pmatrix} \lambda^2 & \lambda(-\varepsilon + 1) \\ \lambda & -\lambda^2 + \varepsilon \end{pmatrix}$$

$\det P(\lambda) = \frac{1}{4}\lambda^2(-\lambda^2 + 2\varepsilon - 1)$ has two zeroes inside the unit disc (at zero) and two zeroes outside the unit disc (since $|\varepsilon| > 1$).

Outline of the Gohberg-Lerer proof

Introduce from the first column of the inverse of T_n (see (1)) the companion matrix

$$K = \begin{pmatrix} -X_{10}^{(n)}(X_{00}^{(n)})^{-1} & I & & & \\ \cdot & & \cdot & & \\ \cdot & & & \cdot & \\ \cdot & & & & \cdot \\ \cdot & & & & I \\ -X_{n0}^{(n)}(X_{00}^{(n)})^{-1} & & & & \end{pmatrix}.$$

Also introduce

$$\hat{K} = \begin{pmatrix} (X_{00}^{(n)})^{-1} & & & & \\ -X_{10}^{(n)}(X_{00}^{(n)})^{-1} & I & & & \\ \cdot & & \cdot & & \\ \cdot & & & \cdot & \\ \cdot & & & & \cdot \\ -X_{n0}^{(n)}(X_{00}^{(n)})^{-1} & & & & I \end{pmatrix}.$$

Gohberg-Lerer Theorem

Theorem 4. *Given* $X_{00}^{(n)}, X_{10}^{(n)}, \cdots, X_{n0}^{(n)}$ *with* $X_{00}^{(n)}$ *invertible, there exists a hermitian block Toeplitz matrix* T_n *such that* (1) *is satisfied if and only if* $X_{00}^{(n)}$ *is hermitian and the matrix equation*

$$H - K^*HK = \mathrm{diag}\left((X_{00}^{(n)})^{-1}, 0, \cdots, 0\right) \quad (2)$$

is solvable.

In this case any hermitian solution H *of* (2) *generates a hermitian block Toeplitz matrix* T_n *satisfying* (1) *by the formula*

$$T_n = \widehat{K}^* \begin{pmatrix} X_{00}^{(n)} & 0 \\ 0 & H \end{pmatrix} \widehat{K}. \quad (3)$$

Conversely, given a hermitian block Toeplitz matrix $T_n = (t_{j-k})_{j,k=0}^{n}$ *satisfying* (1)*, the matrix* $H = T_{n-1} = (t_{j-k})_{j,k=0}^{n-1}$ *is a hermitian solution of* (2)*, and substituting this solution into* (3) *one obtains the given* T_n*.*

Zeros of $P_n(\lambda)$ are eigenvalues of K and conversely.

Wimmer-Chen result on inertia tells us:

Proposition 1. *Let the pair (C, A) be observable, and let $H = H^*$ solve*

$$H - A^* H A = C^* C$$

Then the inertia of H with respect to the imaginary line is the same as the inertia of A with respect to the unit circle.

Setting $A = K$, $C = \text{diag}\left((X_{00}^{(n)})^{-\frac{1}{2}}, 0, \cdots, 0\right)$ in Proposition 1 one obtains Theorem 2 from Theorem 4.

The indefinite case

In case $X_{00}^{(n)}$ is indefinite, the right hand side in (2) is indefinite as well.

This motivates to look at the general Stein equation

$$H - A^* H A = C^* J C$$

with $J = J^*$ invertible and non-definite.

We shall assume throughout that H is invertible as well.

We cannot expect a straightforward relation between in(H) and in(A); generically any combination is possible. Rather one should consider more complicated objects related to H and A.

General Stein Equation

$$H - A^*HA = C^*JC$$

Here A and H are $n \times n$ matrices C is an $m \times n$ matrix.

We assume $\begin{pmatrix} A \\ C \end{pmatrix}$ is nonsingular (e.g., when (C, A) is observable).

For some $P = P^*$ define on $\mathbb{C}^{n+m}$ two indefinite inner products , given by

$$H_1 = \begin{pmatrix} H & 0 \\ 0 & J - P \end{pmatrix}, \qquad H_2 = \begin{pmatrix} H & 0 \\ 0 & J \end{pmatrix}.$$

Introduce two subspaces $V_1 = \text{Im} \begin{pmatrix} I \\ 0 \end{pmatrix}$ and $V_2 = \text{Im} \begin{pmatrix} A \\ C \end{pmatrix}$, and a map $U_0 : V_1 \to V_2$ defined by $U_0 \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} Ax \\ Cx \end{pmatrix}$.

General Stein Equation continued

Then the general Stein equation rewrites as

$$\begin{aligned} & \langle H_2 U_0 \begin{pmatrix} x \\ 0 \end{pmatrix}, U_0 \begin{pmatrix} x \\ 0 \end{pmatrix} \rangle = \\ = \; & \langle (A^* H A + C^* J C) x, x \rangle = \\ = \; & \langle H_1 \begin{pmatrix} x \\ 0 \end{pmatrix}, \begin{pmatrix} x \\ 0 \end{pmatrix} \rangle. \end{aligned}$$

In other words, U_0 is an H_1–H_2–isometry from V_1 onto V_2.

Witt's Theorem

Theorem 5. *Let* $[\cdot,\cdot]_1$ *and* $[\cdot,\cdot]_2$ *be the two scalar products in* $\mathbb{C}^{n+m}$ *defined by the invertible hermitian matrices* H_1 *and* H_2, *respectively:*

$$[x,y]_1 = \langle H_1 x, y\rangle; \quad [x,y]_2 = \langle H_2 x, y\rangle.$$

Assume that $in H_1 = in H_2$. *Let* $U_0 : V_1 \to V_2$, *where* V_1 *and* V_2 *are subspaces in* $\mathbb{C}^{n+m}$, *be a nonsingular linear transformation that preserves the scalar products*

$$[U_0 x, U_0 y]_2 = [x,y]_1 \quad \text{for every} \quad x,y \in V_1.$$

Then there exists a linear transformation U *on* $\mathbb{C}^{n+m}$ *such that* $[Ux,Uy]_2 = [x,y]_1$ *for every vectors* x *and* y, *and* $Ux = U_0 x$ *for every* $x \in V_1$.

Direct Application

Let P be an arbitrary matrix for which

$$in(J - P) = in(J).$$

Later on we shall take $P > 0$, which is always possible.

Proposition 2. *Given P such that $in(J-P) = in(J)$ there exist matrices X and Y such that*

$$\begin{pmatrix} H & 0 \\ 0 & J \end{pmatrix} - \begin{pmatrix} A^* & C^* \\ X^* & Y^* \end{pmatrix} \begin{pmatrix} H & 0 \\ 0 & J \end{pmatrix} \begin{pmatrix} A & X \\ C & Y \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & P \end{pmatrix}.$$

If in addition the pair (C, A) is observable and P is invertible, then also the pair $\left(\begin{pmatrix} 0 & 0 \\ 0 & P \end{pmatrix}, \begin{pmatrix} A & X \\ C & Y \end{pmatrix} \right)$ is observable.

Construction of X and Y

Proposition 3. *Assume that (C, A) is observable and suppose that A is invertible. Set*

$$J_1 = J + JCA^{-1}H^{-1}A^{-*}C^*J,$$

then $in(J_1) = in(J)$. Let Y be any matrix for which

$$Y^*J_1Y < J.$$

Set

$$X = -H^{-1}A^{-*}C^*JY, \text{ and } P = J - Y^*J_1Y.$$

Then

$$\begin{pmatrix} H & 0 \\ 0 & J \end{pmatrix} - \begin{pmatrix} A^* & C^* \\ X^* & Y^* \end{pmatrix} \begin{pmatrix} H & 0 \\ 0 & J \end{pmatrix} \begin{pmatrix} A & X \\ C & Y \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & P \end{pmatrix},$$

and the pair $\left(\begin{pmatrix} 0 & 0 \\ 0 & P \end{pmatrix}, \begin{pmatrix} A & X \\ C & Y \end{pmatrix} \right)$ is observable.

Hence

$$in \begin{pmatrix} H & 0 \\ 0 & J \end{pmatrix} = \widetilde{in} \begin{pmatrix} A & X \\ C & Y \end{pmatrix}.$$

Back to block Toeplitz

Recall

$$H - K^*HK = \text{diag}\left((X_{00}^{(n)})^{-1}, 0, \cdots, 0\right)$$

Here K is the companion matrix introduced earlier.

Take $A = K$, $C = \begin{pmatrix} I_r & 0 & \cdots & 0 \end{pmatrix}$, and $J = (X_{00}^{(n)})^{-1}$.

Proposition 4. *Choose $P > 0$ such that*

$$in((X_{00}^{(n)})^{-1} - P) = in(X_{00}^{(n)})^{-1}$$

and take matrices X and Y such that

$$\begin{pmatrix} K^* & C^* \\ X^* & Y^* \end{pmatrix} \begin{pmatrix} T_{n-1} & 0 \\ 0 & (X_{00}^{(n)})^{-1} \end{pmatrix} \begin{pmatrix} K & X \\ C & Y \end{pmatrix}$$

$$= \begin{pmatrix} T_{n-1} & 0 \\ 0 & (X_{00}^{(n)})^{-1} - P \end{pmatrix}$$

Then $inT_n = \widetilde{in}\begin{pmatrix} K & X \\ C & Y \end{pmatrix}$.

Assume K invertible, i.e., $X_{n0}^{(n)}$ is invertible. This turns out to be generic.

In this case Y is any matrix satisfying

$$Y^*(X_{0n}^{(n)})^{-*}X_{nn}^{(n)}(X_{0n}^{(n)})^{-1}Y < (X_{00}^{(n)})^{-1}$$

and computation of X gives

$$X = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix} = \begin{pmatrix} X_{1n-1}^{(n-1)} \\ \vdots \\ X_{n-1n-1}^{(n-1)} \end{pmatrix} (X_{n0}^{(n)})^{-*}Y$$

Computing the characteristic polynomial of

$$\lambda I - \begin{pmatrix} K & X \\ C & Y \end{pmatrix}$$

and using the Gohberg-Heinig formulas for T_{n-1}^{-1} in terms of the first and last columns of T_n^{-1} we obtain the main result.

Main Theorem

Theorem 6. *Suppose $X_{0n}^{(n)}$ is invertible, and let Y be any matrix for which*

$$Y^*(X_{0n}^{(n)})^{-*}X_{nn}^{(n)}(X_{0n}^{(n)})^{-1}Y < (X_{00}^{(n)})^{-1}.$$

Let $X_0(\lambda)$ and $X_n(\lambda)$ be the two orthogonal matrix polynomials connected to the first and last column of T_n^{-1}. Define $P(\lambda)$ by

$$P(\lambda) = \lambda X_0(\lambda)(X_{00}^{(n)})^{-1} + X_n(\lambda)(X_{0n}^{(n)})^{-1}Y,$$

then the number of positive, respectively, negative eigenvalues of T_n is equal to the number of zeroes of $P(\lambda)$ inside, respectively, outside the unit disc (multiplicities taken into account).

Remarks

The assumption that $X_{00}^{(n)}$ and $X_{0n}^{(n)}$ are invertible can be made generically.

Consider the special case where $X_{00}^{(n)}$ is positive definite. In this case it is obvious that we can take $Y = 0$. So, in that case $P(\lambda) = \lambda X_0(\lambda)(X_{00}^{(n)})^{-1}$. Thus we see that the number of positive, respectively, negative, eigenvalues of T_n is equal to the number of zeroes of $\lambda X_0(\lambda)$ inside, respectively, outside, the unit disc.

Note that the converse also holds: we can take $Y = 0$ only in case $X_{00}^{(n)}$ is positive definite.