Hermitian Forms and Zeros of a Polynomial

Pranshu Gaba [*]

Indian Institute of Science, Bangalore
gabapranshu@ug.iisc.in

November 30, 2017

Abstract

We looked at the general properties of Hermitian (self-adjoint) matrices, and used the Schur-Cohn theorem to find the number of roots of a polynomial lying within and without the unit circle.

1 Introduction

In this paper we see the properties of Hermitian matrices, which are very useful and interesting. We also see and prove the Schur-Cohn theorem to find the number of roots of a polynomial lying within the unit circle.

There are many ways to locate the roots of a polynomial. The Schur-Cohn theorem shows a surprising connection between linear algebra and roots of a polynomial. It will be used to find out how many roots of the polynomial lie inside and outside the unit circle.

First we will define some basic terms that will be used ahead in the paper.

2 Definitions

2.1 Norm of a matrix

2.1.1 Operator norm

Given $A \in \mathbb{M}_n$, define $\|A\| = \sup\limits_{x \neq 0} \dfrac{\|Ax\|}{\|x\|} = \sup\limits_{\|x\|=1} \|Ax\|$ to be the operator norm of A. The triangle inequality $\|A + B\| \leq \|A\| + \|B\|$ is satisfied.

[*] :)

2.1.2 Hilbert-Schmidt norm

The Hilbert-Schmidt norm of matrix A, is defined as the square root of sum of squares of all entries in A.

$$\|A\|_2 = \left(\sum_{i,j} |a_{ij}|^2\right)^{1/2}$$

The operator norm is always less than or equal to the Hilbert-Schmidt norm.

2.2 Inner product

The inner product is a binary operator on two vectors $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$. satisfies the following conditions for all $x, y, z \in V$ and $a \in \mathbb{C}$:

- It is linear in the first term.

  $\langle ax, y\rangle = a\langle x, y\rangle$

  $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$

- It becomes its complex conjugate when the arguments are reversed.

  $\langle x, y\rangle = \overline{\langle y, x\rangle}$

- Inner product of a vector with itself is non-negative.

  $\langle x, x\rangle \geq 0$. Here equality is achieved if and only if $x = 0$.

2.3 Adjoint

Let $A \in \mathbb{M}_n(\mathbb{C})$, a $n \times n$ square matrix with complex entries. The adjoint of matrix A, denoted by A^*, is the matrix that satisfies $\langle A^*x, y\rangle = \langle x, Ay\rangle$.

We see, from the properties of inner product, that the adjoint of a matrix is obtained by taking its transpose, followed by taking the complex conjugate of every element.

If the ij^{th} of A is a_{ij}, then the ij^{th} entry of A^* is $\overline{a_{ji}}$. It follows that $(A^*)^* = A$.

Note that A^*, like A, represents a linear transformation on $\mathbb{C}^n$.

a_{ij} is defined as $\langle Ae_j, e_i\rangle$.

2.4 Positive Definite Matrix

Definition. A matrix $A \in \mathbb{M}_n$ that satisfies $\langle Ax, x\rangle \geq 0$ for all $x \in \mathbb{C}^n$ is called a *positive semidefinite matrix*. If the inequality is strict, $\langle Ax, x\rangle > 0$, then A is called a *positive definite matrix*.

Theorem 1. *All eigenvalues of a positive semidefinite matrix are non-negative.*

Proof. Left to the reader. □

Lemma. *If* $A \geq 0$, *then* $\|A\| = \sup_{\|x\|=1} \langle Ax, x\rangle$

Proof. Left to the reader. □

2.5 Unitary Matrices

A square matrix U is a unitary matrix if $U^*U = I$. The determinant of a unitary matrix is 1. It preserves inner product, $\langle Ux, Uy\rangle = \langle x, y\rangle$.

Proof. Left to the reader □

2.6 Trace

The trace of a matrix is the sum of the diagonal elements of the matrix.
$\mathrm{Tr}(A) = \sum_{i=1}^{n} \langle Ae_i, e_i\rangle$

Theorem 2. *The trace of A^*A is equal to the Hilbert-Schmidt norm of A.* $\mathrm{Tr}(A^*A) = \|A\|_2$

Proof. Left to the reader □

2.7 Hermitian Matrix

A matrix that satisfies $A = A^*$ is called a Hermitian matrix (also known as self-adjoint matrix).

Theorem 3. *All eigenvalues of a Hermitian matrix are real.*

Proof. Let v be an eigenvector of a Hermitian matrix A, and let λ be the corresponding eigenvalue. Then $Av = \lambda v$.

$\langle Av, v\rangle = \langle \lambda v, v\rangle = \lambda\langle v, v\rangle$. Also $\langle Av, v\rangle = \langle v, A^*v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle$

This means $\lambda\langle v, v\rangle = \overline{\lambda}\langle v, v\rangle$ for any v. Since λ is the same conjugate as its complex conjugate, it implies λ is real. □

The converse of this is also true.

Theorem 4 (Converse of Theorem 3)**.** *If $A \in \mathbb{M}_n(\mathbb{C})$ and $\langle Ax, x\rangle \in \mathbb{R}$ for every x, then $A = A^*$.*

Proof. Let $\alpha \in \mathbb{C}$ and $h, g \in \mathbb{C}^n$. Then $\langle A(h + \alpha g), h + \alpha g\rangle = \langle Ah, h\rangle + \alpha\langle Ag, h\rangle + \overline{\alpha}\langle Ah, g\rangle + |\alpha|^2 \langle Ag, g\rangle$

So $\alpha\langle Ag, h\rangle + \overline{\alpha}\langle Ah, g\rangle = \overline{\alpha}\langle h, Ag\rangle + \alpha\langle g, Ah\rangle$
When $\alpha = 1$, $\langle Ag, h\rangle + \langle Ah, g\rangle = \langle h, Ag\rangle + \langle g, Ah\rangle$
When $\alpha = i$, $i\langle Ag, h\rangle - i\langle Ah, g\rangle = -i\langle h, Ag\rangle + i\langle g, Ah\rangle$
$2i\langle Ag, h\rangle = 2i\langle g, Ah\rangle$ ot $\langle Ag, h\rangle = \langle g, Ah\rangle = \langle A^*g, h\rangle$
$Ag = A^*g$ for all g, therefore $A = A^*$. A is Hermitian. □

Corollary. *Every positive semidefinite matrix is Hermitian.*

Proof. Left to the reader. □

The converse is also true.

Theorem 5. *A^*A is always positive semidefinite.*

Proof. $\langle A^*Ax, x\rangle = \langle Ax, Ax\rangle = \|Ax\|^2 \geq 0$ □

Lemma. *If A is Hermitian, then* $\|A\| = \sup_{\|h\|=1} |\langle Ah, h\rangle|$.

Proof. Left to the reader □

Lemma. *If $\langle Ah, h\rangle = 0$ for all h, then $A = 0$.*

Proof. Left to the reader □

2.8 Diagonalization

Hermitian matrices can be diagonalized. For every Hermitian matrix A, there exists a diagonal matrix Λ such that $A = U^*\Lambda U$. Here U is some unitary matrix.

2.9 Projectors

A matrix P is a projector if $P^2 = P$ and $P^* = P$

Theorem 6. *There exists a subspace M of $\mathbb{C}^n$ such that $Pm = m \ \forall m \in M$, and $Px = 0 \ \forall x \in m^\perp$*

Proof. Left to the reader □

2.10 Shift Matrix

Let S, the *shift matrix*, be the $n \times n$ square matrix $\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$.

Note that S is a nilpotent matrix of order n, i.e. S^n is a zero matrix.

Related useful matrices: S^*, SS^*, $I - SS^*$. The last one is a projector.

3 Schur-Cohn Theorem

Given any polynomial $p(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n$ with complex coefficients, we are interested in finding how many of its roots lie within the unit circle and how many roots lie outside. Without loss of generality, let $a_0 = 1$ as it does not change the roots of the polynomial.

Suppose p has roots α_i. Then $p(z) = (z - \alpha_1)(z - \alpha_2)\cdots(z - \alpha_n)$.

Then $p(S)$ is $\begin{bmatrix} a_n & a_{n-1} & \ddots & \ddots & a_1 \\ 0 & a_n & a_{n-1} & \ddots & \ddots \\ 0 & 0 & a_n & \ddots & \ddots \\ 0 & 0 & 0 & \ddots & a_{n-1} \\ 0 & 0 & 0 & 0 & a_n \end{bmatrix}$.

This can be factorized as $p(S) = (S - \alpha_1 I)(S - \alpha_2 I)\cdots(S - \alpha_n I)$. Let $B_j = S - \alpha_j I$.

Next, define q as the polynomial $\overline{a_n} z^n + \overline{a_{n-1}} z^{n-1} + \cdots + \overline{a_0}$. Note that its roots are $\frac{1}{\overline{\alpha_i}}$. We get $q(z) = (1 - \overline{\alpha_1} z)(1 - \overline{\alpha_2} z)\cdots(1 - \overline{\alpha_n z})$. Also, $q(S) = (I - \overline{\alpha_1} S)(I - \overline{\alpha_2} S)\cdots(I - \overline{\alpha_n} S)$. Let $C_j = I - \overline{\alpha_j} S$.

Let H be equal to $\|q(S)x\|^2 - \|p(S)x\|^2$

H can also be written as $\langle (q(S)^* q(S) - p(S)^* p(S))x, x \rangle$.

We can now state the Schur-Cohn theorem:

Theorem 7. *The polynomial p, it will have k roots inside the circle, and $n - k$ roots outside the circle iff k eigenvalues of H are positive and $n - k$ are negative.*

4 Proof

We will first prove the Schur-Cohn theorem for $n = 1$, that is for linear polynomials. It will then be extended to polynomials of higher degrees with the help of the Spectral theorem and the Courant-Fischer theorem.

4.1 Linear Polynomial

Let's write $q(S)$ and $p(S)$ as a product of the linear terms. $q(S)^*q(S) - p(S)^*p(S)$
$= (C_1C_2C_3 \dots C_n)^*(C_1C_2C_3 \dots C_n) - (B_1B_2B_3 \dots B_n)^*(B_1B_2B_3 \dots B_n)$

Let's look at $C_1^*C_1 - B_1^*B_1$ first. Substituting the values of C_1 and B_1, we get

$$\begin{aligned}
&C_1^*C_1 - B_1^*B_1 \\
&= (I - \overline{\alpha_1}S)^*(I - \overline{\alpha_1}S) - (S - \alpha_1 I)^*(S - \alpha_1 I) \\
&= (I - \alpha_1 S^*)(I - \overline{\alpha_1}S) - (S^* - \overline{\alpha_1}I)(S - \alpha_1 I) \\
&= (I - \alpha_1 S^* - \overline{\alpha_1}S + |\alpha_1|^2 S^*S) - (S^*S - \alpha_1 S^* - \overline{\alpha_1}S + |\alpha_1|^2 I) \\
&= I - |\alpha_1|^2 I - S^*S + |\alpha_1|^2 S^*S \\
&= (1 - |\alpha_1|^2)(I - S^*S)
\end{aligned}$$

Note that $I - S^*S$ is a posititve definite matrix. If $|\alpha| < 1$, then the root of the linear polynomial lies within the unit circle. Also note that H has one negative eigenvalue. Similarly, if $|\alpha| > 1$, then the root of the linear polynomial lies outside the unit circle, and the eigenvalue of H is positive. This shows that the Schur-Cohn theorem is true for $n = 1$. We will now extend the proof for all n.

4.2 Spectral theorem

Theorem 8. *Let $A \in \mathbb{M}_n$ be a Hermitian matrix with eigenvalues $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots \leq \lambda_n$. Then A can be written as $U\Lambda U^*$, where U is a unitary matrix, and Λ is a diagonal matrix with real entries.*

Proof. Left to the reader. □

4.3 Courant-Fischer theorem

Theorem 9. *Let $A \in \mathbb{M}_n$ be a Hermitian matrix with eigenvalues $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots \leq \lambda_n$. Then*

$$\lambda_k = \min_{\omega_1,\dots,\omega_{n-k} \in \mathbb{C}^n} \max_{\substack{x \neq 0, x \in \mathbb{C}^n \\ x \perp \omega_1,\dots,\omega_{n-k}}} \frac{\langle Ax, x\rangle}{\langle x, x\rangle}$$

Proof. If $x \neq 0$, then $\frac{\langle Ax,x\rangle}{\langle x,x\rangle} = \frac{\langle U\Lambda U^*x,x\rangle}{\langle U^*x,U^*x\rangle}$
$= \frac{\langle \Lambda U^*x,U^*x\rangle}{\langle U^*x,U^*x\rangle}$. and $\{U^*x : x \neq 0\} = \{x \in \mathbb{C}^n : x \neq 0\}$

Thus if $\omega_1, \dots, \omega_{n-k}$ are given, then

$$\sup_{\substack{x \neq 0 \\ x \perp \omega_1,\dots,\ \omega_{n-k}}} \frac{\langle Ax, x\rangle}{\langle x, x\rangle} = \sup_{\substack{y \neq 0 \\ y \perp U^*\omega_1,\dots,\ U^*\omega_{n-k}}} \frac{\langle \Lambda y, y\rangle}{\langle y, y\rangle}$$

$x \perp \omega$ if and only if $y \perp U^*\omega$.

$$= \sup_{\substack{\langle y,y\rangle=1 \\ y\perp U^*\omega_1,\ \ldots,\ U^*\omega_{n-k}}} \sum_{i=1}^{n} \lambda_i \left|y_i\right|$$

$$\geq \sup_{\substack{\langle y,y\rangle=1 \\ y\perp U^*\omega_1,\ \ldots,\ U^*\omega_{n-k} \\ y_1=y_2=\cdots=y_k-1=0}} \sum_{i=1}^{n} \lambda_i \left|x_i\right|^2$$

$$= \sup_{\substack{\langle y,y\rangle=1 \\ y\perp U^*\omega_1,\ \ldots,\ U^*\omega_{n-k} \\ y_1=y_2=\cdots=y_k-1=0}} \sum_{i=k}^{n} \lambda_i \left|y_i\right|^2$$

$$\geq \lambda_k$$

Let $\omega_1 = x_n, \ldots,\ \omega_{n-k} = x_{k+1}$
If $x \perp \omega_i$, as above, then $x = \sum_{i=1}^{k} c_i x_i$.
$\langle Ax, x\rangle = \langle A \sum_{i=1}^{n} c_i X_i, \sum_{i=1}^{n} c_i x_i\rangle$
$= \langle \sum_{i=1}^{n} c_i \lambda_i x_i, \sum_{i=1}^{n} c_i x_i\rangle$
$= \sum_{i=1}^{k} \lambda_i \left|c_i\right|^2$
$\leq \lambda_k \sum_{i=1}^{k} \left|c_i\right|^2$

□

5 Extensions

5.1 Arbitrary radius

5.2 Limitations

6 Conclusion

Thank You!

7 Acknowledgements

8 References