Variant of Schur's Inequality

Michael Taylor

Let A be a complex $n \times n$ matrix; we write $A \in M(n, \mathbb{C})$. A theorem of Schur implies one can write

$$A = D + N, \tag{1}$$

where, in some orthonormal basis, D is diagonal and N is strictly upper triangular. The diagonal entries of D are the eigenvalues λ_k of A, repeated according to multiplicity, so

$$\sum |\lambda_k|^2 = \|D\|_{\mathrm{HS}}^2 = \|A\|_{\mathrm{HS}}^2 - \|N\|_{\mathrm{HS}}^2. \tag{2}$$

Here $\|A\|_{\mathrm{HS}}^2 = \operatorname{Tr}(A^* A)$ is the square Hilbert-Schmidt norm of A. In particular, we have

$$\sum |\lambda_k|^2 \le \|A\|_{\mathrm{HS}}^2, \tag{3}$$

a result known as Schur's inequality.

As noted in [D], this can be applied to estimate the roots λ_k of a monic polynomial $z^n + a_{n-1} z^{n-1} + \cdots + a_0$, since these roots coincide with the eigenvalues of the companion matrix

$$A = \begin{pmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{n-2} & -a_{n-1} \end{pmatrix}. \tag{4}$$

We obtain from (3) that

$$\sum |\lambda_k|^2 \le \sum_{j=0}^{n-1} |a_j|^2 + (n-1). \tag{5}$$

Note that in going from (2) to (3) you lose something, namely $\|N\|_{\mathrm{HS}}^2$. Now one has $N = 0$ if and only if A is normal, i.e., if and only if $A^* A = A A^*$. The matrices of the form (4) are far from normal. Our goal here is to estimate $\|N\|_{\mathrm{HS}}$ from below in terms of $[A^*, A]$ and improve (3). We will establish the following.

Proposition 1. *If $\{\lambda_k : 1 \le k \le n\}$ are the eigenvalues of $A \in M(n, \mathbb{C})$, counted with multiplicity, then*

$$\sum |\lambda_k|^2 \le \frac{\|A\|_{\mathrm{HS}}^2}{1 + \varphi(\xi(A))^2}, \tag{6}$$

where

$$\xi(A) = \frac{\|[A^*, A]\|_{\text{HS}}}{2\|A\|^2_{\text{HS}}}, \quad \varphi(x) = (1+x)^{1/2} - 1. \tag{7}$$

In light of the elementary estimate

$$\|XY\|_{\text{HS}} \le \|X\|_{\text{HS}} \|Y\|_{\text{HS}}, \tag{8}$$

we have

$$0 \le \xi(A) \le 1 \tag{9}$$

for all nonzero $A \in M(n, \mathbb{C})$. Note that $\varphi(x)$ is smooth and monotonically increasing in $x \in [0, 1]$, with

$$\varphi(0) = 0, \quad \varphi(1) = \sqrt{2} - 1 \approx 0.414. \tag{10}$$

Wanting to estimate $\|N\|_{\text{HS}}$ from below, we proceed to estimate $\|[A^*, A]\|_{\text{HS}}$ from above. Note that

$$[A^*, A] = [N^*, D] + [\overline{D}, N] + [N^*, N]. \tag{11}$$

Using (8) and the triangle inequality $\|X + Y\|_{\text{HS}} \le \|X\|_{\text{HS}} + \|Y\|_{\text{HS}}$, we obtain

$$\begin{aligned} \|[A^*, A]\|_{\text{HS}} &\le 2\|[\overline{D}, N]\|_{\text{HS}} + \|[N^*, N]\|_{\text{HS}} \\ &\le 4\|D\|_{\text{HS}} \|N\|_{\text{HS}} + 2\|N\|^2_{\text{HS}}, \end{aligned} \tag{12}$$

so

$$\|N\|^2_{\text{HS}} + 2\|D\|_{\text{HS}} \|N\|_{\text{HS}} \ge \frac{1}{2} \|[A^*, A]\|_{\text{HS}}. \tag{13}$$

Completing the square on the left side of (13) gives

$$\begin{aligned} \|N\|_{\text{HS}} &\ge \left(\frac{1}{2}\|[A^*, A]\|_{\text{HS}} + \|D\|^2_{\text{HS}}\right)^{1/2} - \|D\|_{\text{HS}} \\ &= \|D\|_{\text{HS}} \left[\left(1 + \frac{\|[A^*, A]\|_{\text{HS}}}{2\|D\|^2_{\text{HS}}}\right)^{1/2} - 1\right] \\ &\ge \|D\|_{\text{HS}}\, \varphi(\xi(A)), \end{aligned} \tag{14}$$

with φ and $\xi(A)$ as in (7), the last inequality holding because $\|D\|^2_{\text{HS}} \le \|A\|^2_{\text{HS}}$. Recalling (2), we have

$$\|D\|^2_{\text{HS}} \le \|A\|^2_{\text{HS}} - \|D\|^2_{\text{HS}}\, \varphi(\xi(A))^2, \tag{15}$$

which gives the asserted estimate (6).

Remark. One loses something in passing to the last inequality in (14). If we stop before that, we obtain

$$\|N\|_{\text{HS}} \ge \|D\|_{\text{HS}}\, \varphi\Big(\frac{K}{2\|D\|^2_{\text{HS}}}\Big), \quad K = \|[A^*, A]\|_{\text{HS}}, \tag{16}$$

which leads to

$$\left(1 + \varphi\Big(\frac{K}{2\|D\|^2_{\text{HS}}}\Big)^2\right) \|D\|^2_{\text{HS}} \le \|A\|^2_{\text{HS}}. \tag{17}$$

As an improvement over (3), this is sharper than (6)–(7), though less explicit.

Reference

[D] E. Deutsch, Solution II to Problem #11008, American Math. Monthly 112 (2005), p. 92.