Mathematical
Inequalities
& Applications
Volume 13, Number 3 (2010), 625–628

SCHUR'S INEQUALITY FOR THE DIMENSION OF COMMUTING FAMILIES OF MATRICES

N. S. KARAMZADEH

(*Communicated by G. Styan*)

Abstract. We present an elementary proof for the above inequality. This in particular leads us easily to see that the maximal linearly independent sets of commuting $n \times n$ matrices, for $n \geqslant 4$, necessarily consist of either nilpotent elements or units.

1. Introduction

In [7], Schur proved that the cardinality of any linearly independent set of commuting $n \times n$ complex matrices is less than or equal to $g(n) = \lfloor \frac{n^2}{4} \rfloor + 1$. He also determined linearly independent sets with cardinality $g(n)$, consisting of commuting matrices. Later, Jacobson [4] proved the same inequality for any field K. Using representation theoretic methods, Gustafson in [2] gives an elegant proof of this inequality and by applying the same methods he also finds a lower bound for the dimensions of maximal commutative subalgebras of $n \times n$ matrices over a field K. Some authors have proved Schur's inequality by explicit manipulation of matrices, see [8], [9] and for a more recent one see [5]. Using the methods in [2] and in answering a question in [2], Cowsik [1] gives an upper bound for the length of a commutative Artinian ring in terms of the length of a faithful module and automatically provides a short proof for a generalization of Schur's inequality. Our aim in this short note is to give a very simple proof of this inequality. But before presenting the proof, let us give some reasons for our claim of its simplicity. Suppose that S is a maximal linearly independent set of commuting linear transformations on a finite-dimensional vector space V over a field K. Let I be the identity transformation and $R = K[S \cup \{I\}]$ be the algebra generated by $S \cup \{I\}$ over K. R is a commutative K-subalgebra of $\mathrm{Hom}_K(V,V)$ and Schur's inequality asserts that $\dim_K R \leqslant g(\dim_K V)$. Generally, in proofs of this inequality which are not by manipulation of matrices, the authors first consider the case that R is a local ring and then go to the general case. In order to be allowed to do so, they usually apply the result which says every commutative Artinian ring R is a direct product of Artinian local rings and also decompose each faithful finitely generated R-module as a direct sum of submodules which are faithful over local components of R, see for example [1],

Mathematics subject classification (2010): 15A04, 15A03.
Keywords and phrases: Schur's inequality, commutativity.

© ELEMENT, Zagreb
Paper MIA-13-43

[2], for two clever proofs of this kind. But in our proof, by just using some elementary facts from linear algebra we are left with the case that R is forced to be local and we make no use of its Artinian property. Finally, as a consequence of this proof, we trivially observe that if S is a maximal linearly independent set of commuting $n \times n$ matrices, where $n \geqslant 4$, then its elements must be either nilpotent or units. The latter fact can also be obtained essentially, from Theorem 2, Lemma 3 and Theorem 3 in [4].

2. Schur's inequality revisited

We need the following trivial fact, see also [2], [4].

LEMMA. *Let* $g(n) = \lfloor \frac{n^2}{4} \rfloor + 1$ *and* $n = d_1 + d_2 + ... + d_k$, *where each* d_i *is a positive integer. Then* $g(n) \geqslant g(d_1) + g(d_2) + ... + g(d_k)$.

Proof. An easy induction on k. □

THEOREM. (Schur) *Let* S *be a set of mutually commuting linearly independent matrices of order* n *over a field* K. *Then* $|\, S \,| \leqslant g(n)$.

Proof. First we recall that if K is a subfield of a field L, then whenever a system of homogeneous linear equations with coefficients in K, has a nontrivial solution in L, it has also a nontrivial solution in K. This immediately implies that S is also linearly independent over L. Thus without loss of generality one can assume that K is algebraically closed and S is maximal with respect to the above property. Let $V = K^n$ and $R = K[S \cup \{I\}] \subseteq \mathrm{Hom}_K(V,V)$. Clearly, R is a maximal commutative subalgebra of $\mathrm{Hom}_K(V,V)$. We proceed by induction on n. For $n = 1$ the claim is trivial. Suppose that the theorem is true for matrices of order less than n. First, let us assume that an element $A \in S$ has at least two distinct characteristic values and $P(x) = (x-\lambda_1)^{r_1}(x-\lambda_2)^{r_2}...(x-\lambda_k)^{r_k}$ be its minimal polynomial, where we assume the λ_i distinct. Then we can write $V = W_1 \oplus W_2 \oplus ... \oplus W_k$, where each W_i is the null space of $(A-\lambda_i I)^{r_i}$. Clearly, each W_i is invariant under every element of R. Now for each element $x \in V$ we have the unique representation $x = w_1 + ... + w_k$ with $w_i \in W_i$. Given $B \in R$ define $B_i = B_i(B)$ by $B_i(x) = B(w_i)$. If $C \in R$ the fact $C(W_l) \subseteq W_l$ for all $l = 1,...,k$ implies for all admissible $i \neq j$, $B_iC\,|_{W_i} = CB_i\,|_{W_i}$ and $B_iC\,|_{W_j} = 0 = CB_i\,|_{W_j}$. So $B_i \in R$. Letting $R_i =$ the span of all $B_i(B)$ as B ranges over R, we have $R = \oplus_{i=1}^{k} R_i$. So by induction hypothesis, $\dim_K(R) = \sum_1^k \dim_K(R_i) \leqslant \sum_1^k g(\dim_K(W_i)) \leqslant g(n)$ and we are done. Therefore we may assume that no element of S has two distinct characteristic values. This means that for each $A \in S$ there exists $r_a \in K$ such that $A - r_a I$ is nilpotent. But each element T of R is of the form $T = \lambda_0 I + \sum_{A_j \in S} \lambda_j A_j$, where λ_0, $\lambda_j \in K$. Hence T may be written as $T = \mu_0 I + \sum_{A_j \in S} \lambda_j (A_j - r_j I)$, where $\mu_0 \in K$ and $A_j - r_j I$ is nilpotent. Clearly $\sum \lambda_j (A_j - r_j I)$, as a finite sum of commuting nilpotent elements, is nilpotent. Consequently, T is either nilpotent or a unit element of R. This immediately implies that R is a local ring and in fact its maximal ideal, M say, is its only prime ideal which consists of the nilpotent elements of R. We also note that V is a faithful R-module and it is well-known (by a variant of Nakayama's lemma) that whenever $\{v_1 + MV,\ v_2 + MV, ...,\ v_m + MV\}$ is a basis for $\frac{V}{MV}$ as a vector space over

$\frac{R}{M}$ then $v_1, v_2, ..., v_m$ generate V as an R-module, see for example [6, Proposition 9.3] or [3, p390]. Observe that $\frac{R}{M}$ is a finite extension of K, i.e., $K \cong \frac{R}{M}$ or, more simply, even without invoking that K is algebraically closed, we have $R = KI + M$; hence $\frac{R}{M} \cong K$. Thus $\dim_K MV = \mathrm{dim_K V} - \mathrm{dim_K} \frac{\mathrm{V}}{\mathrm{MV}} = \mathrm{n} - \mathrm{m}$. We also note that for each v_i, $i = 1, ..., m$, Mv_i is a K-subspace of MV and therefore $\dim_K Mv_i \leqslant n - m$. Now consider $Mv_1 \oplus Mv_2 \oplus ... \oplus Mv_m$ as a subspace of $\overbrace{V \oplus V \oplus ... \oplus V}^{m}$ and define the K-linear map $\varphi : M \to Mv_1 \oplus Mv_2 \oplus ... \oplus Mv_m$ by $\varphi(A) = Av_1 + Av_2 + ...Av_m$ for all $A \in M$ and note that we might have $Mv_i = 0$ for some i. Inasmuch as V is a faithful R-module, we infer that φ is injective and therefore $\dim_K M \leqslant m(n-m) \leqslant \frac{n^2}{4}$, i.e., $\dim_K M \leqslant \lfloor \frac{n^2}{4} \rfloor$. But $\dim_K R = 1 + \dim_K M \leqslant 1 + \lfloor \frac{n^2}{4} \rfloor$ and this completes the proof. □

REMARK. In the second part of the previous proof we note that every element of R is of the form $\lambda I + A$, where A is nilpotent. Motivated by this, we see easily that for each $0 \neq \lambda_{ij}$, $0 \neq \lambda$, $\mu_{ij} \in K$, the set $S = \{ \mu_{ij} I + \lambda_{ij} E_{ij} : 1 \leqslant i \leqslant m, m+1 \leqslant j \leqslant n\} \cup \{\lambda I\}$, where $m = \lfloor \frac{n}{2} \rfloor$ and $\{E_{ij}\}$ are matrix units, is a maximal linearly independent set of commuting $n \times n$ matrices over a field K and $\mid S \mid = g(n)$. Therefore $R = K[S]$ is a maximal commutative local subalgebra of $\mathrm{Hom}_K(V,V)$ with $\dim_K R = g(n)$. From this and the simple fact that if $n = d_1 + d_2 \geqslant 4$, where d_1 and d_2 are positive integers, then $g(n) > g(d_1) + g(d_2)$; we infer that no commutative subalgebra R' of $\mathrm{Hom}_K(V,V)$ of dimension $g(n)$ is decomposable, that is to say, R' must be local (note, here we are using the fact that R' is Artinian), see also [2], [4]. Consequently, every maximal set of linearly independent commutative elements in $\mathrm{Hom}_K(V,V)$ consists only of nilpotent elements or units.

Acknowledgements. I am grateful to three anonymous referees for reading this article carefully and giving useful comments. I would also like to thank Professor O. A. S. Karamzadeh who taught me both linear algebra and ring theory.

REFERENCES

[1]] R. C. COWSIK, *A short note on the Schur-Jacobson theorem,* Proc. Amer. Math. Soc., **2**, 118 (1993), 675–676.
[2] W. H. GUSTAFSON, *On maximal commutative algebras of linear transformation,* J. Algebra, **1**, 42 (1976), 557–563.
[3] T. HUNGERFORD, *Algebra,* GTM 73, Springer, 1974.
[4] N. JACOBSON, *Schur's theorems on commutative matrices*, Bull. Amer. Math. Soc., **50** (1946), 431–436.
[5] M. MIRZAKHANI, *A simple proof of a theorem of Schur,* Amer. Math. Monthly, **3**, 105 (1998), 260–261.
[6] R. Y. SHARP, *Steps in Commutative Algebra,* Cambridge university press, 1990.
[7] I. SCHUR, *Zur theorie vertauschbaren matrizen,* Crelle's J., **130** (1905), 66–76.
[8] D. A. SUPRUNENKO AND R. I. TYSHKEVICH, *Perestanovochnyye Matrisy*, Nauk and Tekhnika press, Minsk, 1966. English translation: *Commutative Matrices*, Academic press, New York, 1968.

[9] WAN ZHE-XIAN AND LI GEN-DAO, *The two theorems of Schur on commutative matrices,* Chinese Math., **5** (1964), 156–165.

(Received July 17, 2007)

N. S. Karamzadeh
Department of Mathematics
Faculty of Mathematical Science
Shahid Beheshti University
Tehran, Iran
e-mail: n-shahni@cc.sbu.ac.ir