On the hermitian normalform of a matrix and Sylvester's law of nullity.

By E. EGERVÁRY in Budapest.

1. It is known[1]) that any square matrix **A** can be decomposed into the product of a non-singular matrix **T** and of a matrix **H** which has the hermitian normalform characterized by the following properties

α) is triangular, i. e. it has the form ◥ or ◣,

β) all the diagonal elements a_{ii} satisfy $a_{ii}^2 - a_{ii} = 0$ i. e. $a_{ii} = 1$ or 0,

γ) each row of **H** (or of **H***)[2]) whose diagonal element is zero contains only 0 elements,

δ) each column of **H** (or of **H***) whose diagonal element is 1 contains (except this 1) only 0 elements.

Throughout this paper the term *quasi-hermitian* matrix will be used to denote a matrix which satisfies the conditions α) β) γ) only.

With regard to the following considerations it is very essential to remark that the given matrix **A** can be decomposed in each of the following manners

$$\left.\begin{array}{ll} \mathbf{A} = \mathbf{T}◥; & \mathbf{A} = ◥\mathbf{T} \\ \mathbf{A} = \mathbf{T}◣; & \mathbf{A} = ◣\mathbf{T}. \end{array}\right\} \tag{1}$$

First of all we shall indicate a method for the hermitian (or quasihermitian) decomposition which seems to be simpler then those published hitherto[3]).

In order to get a decomposition of the form $\mathbf{A} = \mathbf{T}◥$ we use te identity

$$\mathbf{A} = \begin{bmatrix} a_{11} & & \\ & \cdot & \\ & & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} \left[1, \frac{a_{\nu 2}}{a_{\nu 1}}, \ldots, \frac{a_{\nu n}}{a_{\nu 1}}\right] + \left[\begin{array}{c|c} 0 & \mathbf{A}_1' \\ \vdots & \\ 0 & 0 \cdots 0 \\ \vdots & \\ 0 & \mathbf{A}_2' \end{array}\right]; \qquad \begin{bmatrix} \mathbf{A}_1' \\ \mathbf{A}_2' \end{bmatrix} = \mathbf{A}'$$

if the first column of **A** contains at least one non-zero element $a_{\nu 1}$, or the

[1]) See e. g. C. C. MAC DUFFE, Vectors and Matrices, Carus Mathematical Monographs, 7 (1943).

[2]) **H*** is the transpose of **H**.

[3]) See also E. EGERVÁRY, On a property of the projector matrices and its application to the canonical representation of matrix functions. *Acta Sci. Math. (Szeged)*, **15** (1953), pp. 1—6.

identity

$$\mathbf{A}=\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}[0,0,\ldots,0]+\left[\begin{array}{c|c}0 & \\ 0 & \\ \vdots & \mathbf{A}'\\ 0 & \end{array}\right]\qquad (\alpha_i \text{ arbitrary})$$

if all the elements of the first column are 0.

Treating in the same way $\mathbf{A}'$ and so on, one arrives obviously to a decomposition

$$\mathbf{A}=\mathbf{B}◥$$

where in case of convenient choice of the elements α_i $\mathbf{B}$ is non-singular and ◥ has the quasihermitian form.

Finally premultiplication of ◥ by a convenient non-singular (unimodular) matrix reduces all the elements above the non-vanishing diagonalelements to 0 and thus the hermitian normalform is atteined.

Starting with the last column of $\mathbf{A}$ one arrives to $\mathbf{A}=\mathbf{T}$◤ $=\mathbf{T}_1$◣ and the decomposition of $\mathbf{A}^*=\mathbf{T}$◥ yields $\mathbf{A}=$ ◥$^*\mathbf{T}^*=$ ◣$\mathbf{T}^*$.

2. The main object of this paper is to show that the hermitian (as well as the quasihermitian) normalform of a matrix can be advantageously used to the straightforward proof of SYLVESTER's law of nullity which states that

The nullity of the product of two matrices is at least as great as the nullity of either factor, and at most as great as the sum of the nullity of the factors.

The nullity of a square matrix $\mathbf{A}$ is defined as the order of $\mathbf{A}$ — rank of $\mathbf{A}$. Hence using the designation $\varrho(\mathbf{A})$ for the rank of an n-th order matrix, SYLVESTER's law of nullity can be expressed by the following inequalities

$$\varrho(\mathbf{A})+\varrho(\mathbf{B})-n\leqq\varrho(\mathbf{AB})\leqq\min(\varrho(\mathbf{A}),\varrho(\mathbf{B})). \tag{2}$$

The inequality on the right is an obvious consequence of the definition of the rank and it implies immediately the well-known fact that multiplication by a non-singular matrix does not change the rank of a matrix.

The inequality on the left constitutes the essential content of SYLVESTER's law of nullity and for this inequality there exists — as far as the author is informed — no straightforward, elementary proof. There are proofs which consider the number of the distinct solutions of the equation $\mathbf{AB}x=0$[4]), while other proofs[5]) are based on theorems concerning the affin transformations of the n-dimensional vector-space.

We prove first SYLVESTER's law of nullity for two diagonal matrices $\langle a_{11},a_{22}\ldots a_{nn}\rangle,\langle b_{11},b_{22}\ldots b_{nn}\rangle$ whose elements satisfy

$$a_{ii}^2-a_{ii}=0,\quad b_{ii}^2-b_{ii}=0.$$

---

[4]) See f. i. G. A. DIRAC, Sylvester's law of nullity, *Math. Gazette*, **34** (1950), p. 305.

[5]) See f. i. Г. Е. Шилов, введение в теорию линейных пространств, 1952, pp. 118—119.

In this case we have obviously

$$\varrho\langle a_{ii}\rangle=\sum_1^n a_{ii},\ \varrho\langle b_{ii}\rangle=\sum_1^n b_{ii}$$

and

$$\varrho(\langle a_{ii}\rangle\langle b_{ii}\rangle)=\varrho(\langle a_{ii}b_{ii}\rangle)=\sum_1^n a_{ii}b_{ii}.$$

But it is clear that $1-a_{ii}\geqq 0,\ 1-b_{ii}\geqq 0$, hence

$$\sum_1^n(1-a_{ii})(1-b_{ii})=n-\sum_1^n a_{ii}-\sum_1^n b_{ii}+\sum_1^n a_{ii}b_{ii}\geqq 0$$

$$\varrho(\langle a_{ii}\rangle\langle b_{ii}\rangle)=\Sigma a_{ii}b_{ii}\geqq\Sigma a_{ii}+\Sigma b_{ii}-n=\varrho(\langle a_{ii}\rangle)+\varrho(\langle b_{ii}\rangle)-n,$$

thus our proof is complete.

Let us now consider two equally situated triangular matrices $◥_a$ and $◥_b$ whose diagonal elements satisfy $a_{ii}^2-a_{ii}=0,\ b_{ii}^2-b_{ii}=0$. In this case we have obviously

$$\varrho(◥_a ◥_b)\geqq\Sigma a_{ii}b_{ii}\geqq\Sigma a_{ii}+\Sigma b_{ii}-n$$

and if $◥_a$ and $◥_b$ are hermitian (or quasihermitian) normalforms then

$$\varrho(◥_a)=\Sigma a_{ii},\quad \varrho(◥_b)=\Sigma b_{ii},$$

consequently we have proved the following special case of SYLVESTER's law of nullity:

If $◥_a$ and $◥_b$ are (equally situated) hermitian or quasihermitian normalforms, then

(3) $$\varrho(◥_a ◥_b)\geqq\varrho(◥_a)+\varrho(◥_b)-n.$$

Suppose now that **A** and **B** are two arbitrary n-th order matrices and apply the decompositions

$$\mathbf{A}=\mathbf{T}_a ◥_a,\ \mathbf{B}=◥_b\mathbf{T}_b$$

where $|\mathbf{T}_a|\neq 0, |\mathbf{T}_b|\neq 0$ and $◥_a, ◥_b$ are hermitian or quasihermitian normalforms. Then we have obviously

$$\varrho(\mathbf{A})=\varrho(◥_a),\ \varrho(\mathbf{B})=\varrho(◥_b)$$

$$\varrho(\mathbf{AB})=\varrho(\mathbf{T}_a ◥_a ◥_b \mathbf{T}_b)=\varrho(◥_a ◥_b)$$

and substitution of these values of $\varrho(◥_a)$, $\varrho(◥_b)$, $\varrho(◥_b ◥_b)$ in (3) immediately gives

$$\varrho(\mathbf{AB})\geqq\varrho(\mathbf{A})+\varrho(\mathbf{B})-n,$$

i. e., (2).

Corollary. If $\mathbf{AB}=0$ then $\varrho(\mathbf{AB})=0$ and in this case it follows from (2) that

If the product of two square matrices **AB** *is null the sum of their ranks cannot exceed their order:*

(4) $$\varrho(\mathbf{A})+\varrho(\mathbf{B})\leqq n.$$

3. It is seen that the proof of SYLVESTER's law of nullity can be based on the quasihermitian decomposition of a matrix which requires considerably simpler operations than the hermitian decomposition.

Nevertheless we took into consideration the hermitian normalform too because it exhibits some remarkable properties to the discussion of which we now proceed.

Let us suppose that the two n-th order matrices $\mathbf{P}_1, \mathbf{P}_2$ satisfy the equations

$$\mathbf{P}_1 + \mathbf{P}_2 = \mathbf{E}, \tag{5}$$

$$\mathbf{P}_1\mathbf{P}_2 = \mathbf{P}_2\mathbf{P}_1 = 0 \tag{6}$$

where $\mathbf{E}$ denotes the n-th order unit-matrix. (5) implies that

$$\varrho(\mathbf{P}_1) + \varrho(\mathbf{P}_2) \geqq \varrho(\mathbf{E}) = n.$$

(6) and the above-mentioned corollary imply

$$\varrho(\mathbf{P}_1) + \varrho(\mathbf{P}_2) \leqq n,$$

hence

$$\varrho(\mathbf{P}_1) + \varrho(\mathbf{P}_2) = n. \tag{7}$$

Furthermore (5) and (6) imply

$$\mathbf{P}_1^2 = \mathbf{P}_1, \quad \mathbf{P}_2^2 = \mathbf{P}_2. \tag{8}$$

The equations (5)—(8) clearly show that $\mathbf{P}_1$ and $\mathbf{P}_2$ are complementary projectors, i. e., the right eigenspace of either is the orthogonal complement of the left eigenspace of the other.

It has been pointed out[6] that the complete solution of a system of homogenous linear equations $\mathbf{Ax} = 0$ is equivalent to the determination of the orthogonal complement of the sub-space spanned by the row vectors of $\mathbf{A}$.

Thus if $\varrho(\mathbf{A}) = r$, the complete solution of $\mathbf{Ax} = 0$ requires the determination of a matrix $\mathbf{X}$ such that

$$\mathbf{AX} = 0, \quad \varrho(\mathbf{X}) = n - r \quad \text{or} \quad \varrho(\mathbf{X}) + \varrho(\mathbf{A}) = n.$$

Suppose now that $\mathbf{A}$ is a projector, i. e. $\mathbf{A}^2 = \mathbf{A}$. We see that in this case the complete solution of $\mathbf{Ax} = 0$, i. e., the orthogonal complement of the sub-space spanned by the row vectors of $\mathbf{A}$ is immediately given by the complementary projector $\mathbf{X} = \mathbf{E} - \mathbf{A}$ because

$$\mathbf{A}(\mathbf{E} - \mathbf{A}) = \mathbf{A} - \mathbf{A}^2 = 0$$

and by (7)

$$\varrho(\mathbf{E} - \mathbf{A}) = n - \varrho(\mathbf{A}) = n - r.$$

We shall prove now that

If a matrix has the hermitian normal form ◥ *then it is a projector, i.e. it satisfies the equation*

$$◥^2 = ◥.$$

[6] See [1].

To this purpose we prove first the folloving

L e m m a. *If* **A** *is an arbitrary matrix and* **P** *is a projector, then the matrix* $\mathbf{Q}=\mathbf{P}+\mathbf{PA}(\mathbf{E}-\mathbf{P})$ *is also a projector.*

Indeed

$$(\mathbf{E}-\mathbf{Q})\mathbf{Q}=(\mathbf{E}-\mathbf{PA})(\mathbf{E}-\mathbf{P})\mathbf{P}(\mathbf{E}+\mathbf{A}(\mathbf{E}-\mathbf{P}))=0$$

since $(\mathbf{E}-\mathbf{P})\mathbf{P}=0$.

Consider now a matrix ◥ in the hermitian normalform. By condition β) all its diagonal elements satisfy $a_{ii}(1-a_{ii})=0$, hence the diagonalmatrix $\langle a_{ii}\rangle$ is obviously a projector. Further, having regard to conditions γ) δ) it is easy to see that

$$\text{◥}=\langle a_{ii}\rangle+[a_{ii}a_{ij}(1-a_{jj})]=\langle a_{ii}\rangle+\langle a_{ii}\rangle\text{◥}\langle 1-a_{ii}\rangle$$

i. e. ◥ is of the type considered in the above lemma, consequently

$$\text{◥}^2=\text{◥} \qquad \text{q. e. d.}$$

According to this result, the complete solution of a system of homogenous linear equations $\mathbf{Ax}=0$ can be obtained in the following way. Decompose **A** into the product **T**◥ of a non-singular matrix **T** and the hermitian normalform ◥. Then $\mathbf{Ax}\equiv\mathbf{T}\text{◥}\mathbf{x}=0$ is equivalent to $\text{◥}\mathbf{x}=0$ and a system of $n-\varrho(\mathbf{A})=n-\varrho(\text{◥})$ distinct solutions of $\mathbf{Ax}=0$ is furnished by the $n-\varrho(\text{◥})$ columns of $\mathbf{E}-\text{◥}$ which are not 0.

4. As an illustration of the method described above we shall find a comlete system of solutions of

$$\mathbf{Ax}=\begin{bmatrix}2&-2&5&1\\1&-1&2&1\\0&0&1&-1\\1&-1&1&2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=0.$$

Starting with the left column of **A** we have

$$\begin{bmatrix}2&-2&5&1\\1&-1&2&1\\0&0&1&-1\\1&-1&1&2\end{bmatrix}=\begin{bmatrix}2\\1\\0\\1\end{bmatrix}\cdot[1\ -1\ \ 2\ \ 1]+\begin{bmatrix}0&0&1&-1\\0&0&0&0\\0&0&1&-1\\0&0&-1&1\end{bmatrix}=$$

$$\begin{bmatrix}2\\1\\0\\1\end{bmatrix}\cdot[1\ -1\ 2\ 1]+\begin{bmatrix}\alpha_1\\\alpha_2\\\alpha_3\\\alpha_4\end{bmatrix}[0\ 0\ 0\ 0]+\begin{bmatrix}1\\0\\1\\-1\end{bmatrix}\cdot[0\ 0\ 1\ -1]+$$

$$+\begin{bmatrix}\beta_1\\\beta_2\\\beta_3\\\beta_4\end{bmatrix}[0\ 0\ 0\ 0]=\begin{bmatrix}2&\alpha_1&1&\beta_1\\1&\alpha_2&0&\beta_2\\0&\alpha_3&1&\beta_3\\1&\alpha_4&-1&\beta_4\end{bmatrix}\cdot\begin{bmatrix}1&-1&2&1\\0&0&0&0\\0&0&1&-1\\0&0&0&0\end{bmatrix}=\mathbf{TK}.$$

Here $\mathbf{K}$ is quasihermitian and the elements α_i, β_j are arbitrary and can be chosen so that $|\mathbf{T}| \neq 0$. Finally permultiplication by a convenient unimodular matrix reduces $\mathbf{K}$ to the hermitian normalform:

$$\begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \text{◥}.$$

Consequently we have $\varrho(\mathbf{A}) = 2$ and

$$\text{◥}(\mathbf{E} - \text{◥}) = \begin{bmatrix} 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 0 & -3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} = 0$$

hence the $n - \varrho(\mathbf{A}) = 2$ columns of $\mathbf{E} - \text{◥}$ which are not 0 constitute a complete system of solutions of $\mathbf{A}\mathbf{x} = 0$, i. e.

$$\mathbf{AX} = \begin{bmatrix} 2 & -2 & 5 & 1 \\ 1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & -3 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} = 0, \ \varrho(\mathbf{X}) = 2.$$

(Received July 2, 1953.)