ON CONDITIONS ON X SUCH THAT XAX* IS HERMITIAN

By

YIK-HOI AU-YEUNG [1]

(Received November 29, 1969)

1. Introduction and definitions. We denote by F the field R of real numbers, the field C of complex numbers, or the skew field H of real quaternions, by $\mathfrak{M}(n, F)$ the set of all $n\times n$ matrices with elements in F and by F^n the n-dimensional left vector space over F. If $A\in\mathfrak{M}(n, F)$, we define $Z_F(A)=\{X\in\mathfrak{M}(n, F): XAX^*=0\}$, rank $Z_F(A)$ $=\max$ {rank $X: X\in Z_F(A)$} and $z_F(A)=\{u\in F^n: uAu^*=0\}$. Here and in what follows we regard u as a $1\times n$ matrix and identify a 1×1 matrix with its single element. For any $A\in\mathfrak{M}(n, F)$, we denote by A^* its conjugate transpose (if $F=R$, then the term "conjugate transpose" merely means "transpose", and A is said to be ***hermitian*** if $A=A^*$, ***skew-hermitian*** if $A=-A^*$ and ***unitary*** if $AA^*=I_n$, where I_n is the $n\times n$ identity matrix (if $F=R$, then the terms "hermitian", "skew-hermitian" and "unitary" mean "symmetric", "skew-symmetric" and "orthogonal" respectively).

Given $A, X\in\mathfrak{M}(n, F)$, let $S=A-A^*$ so that S is skew-hermitian. Then XAX^* is hermitian if and only if $X\in Z_F(S)$. Therefore, the problem of finding condition on X such that XAX^* is hermitian is equivalent to that of finding conditions on X such that $X\in Z_F(S)$.

The purpose of this note is: (i) to prove certain properties of $z_F(B)$ (Theorem 1) and use them to determine the rank $Z_F(S)$ (Theorem 2), and (ii) to derive a simple method by means of which we can obtain a short proof of a result of ***Tihomirov*** [1] (Theorem 3) and a correct and more detailed answer (I hope) of Tihomirov's another problem [2] (Theorem 4) for all cases of F. (In [1] and [2], only the case $F=C$ has been considered.)

2. The rank Z_F(S).

Theorem 1. *Let $B\in\mathfrak{M}(n, F)$ be skew-hermitian (or hermitian) and $\{u_1,\cdots, u_m\}$ a maximal independent set in $z_F(B)$ such that $u_iBu_j^*=0$ for all $i, j=1,\cdots m$. Then*

(i) *the integer m is independent of the choice of the maximal independent set, and*

(ii) $m=$rank $Z_F(B)$.

---

[1] The author wishes to thank Professor *Y. C. Wong* for his advice during the preparation of this note.

Proof. Let $\{v_1, \cdots, v_k\}$ be an independent set of $z_F(B)$ such that $v_p B v_q^* = 0$ for all $p, q = 1, \cdots, k$. Suppose $k > m$, and suppose $v_1, \cdots, v_l \epsilon L = L\{u_1, \cdots, u_m\}$ and $v_{l+1}, \cdots, v_k \notin L$, where $0 \leqslant l \leqslant m$ and $L\{u_1, \cdots, u_m\}$ denotes the subspace spanned by $u_1, \cdots, u_m$ (over F). If $l < m$, then we can decompose L into $L = L\{v_1, \cdots, v_l\} \oplus L_1$. Let $w_{l+1}, \cdots, w_m$ be a basis of L_1. Since $k > m$, there exists $v_0 = \sum_{r=l+1}^{k} \lambda_r v_r (\neq 0)$ in $z_F(B)$ such that $v_0 B w_s^* = 0$ for all $s = l+1, \cdots, m$. Furthermore, since $v_0 B v_t^* = 0$ for all $t = 1, \cdots, l$ and B is skew-hermitian (or hermitian), the maximality of the independent set$\{u_1, \cdots, u_m\}$ requires that $v_0 \epsilon L$. Without loss of generality, we may assume the coefficient λ_{l+1} in v_0 is not zero. Define

$$x_p = \begin{cases} v_p & \text{if } p \neq l+1, \\ v_0 & \text{if } p = l+1, \end{cases} \qquad p = 1, \cdots, k.$$

Then $\{x_1, \cdots, x_k\}$ is an independent set of $z_F(B)$ such that $x_p B x_q^* = 0$ for all $p, q = 1, \cdots, k$, and $x_1, \cdots, x_{l+1} \epsilon L$. By continuing this processes, we can at last find an independent set $\{y_1, \cdots, y_k\}$ of $z_F(B)$ such that $y_p B y_q^* = 0$ for all $p, q = 1, \cdots, k$ and $y_1, \cdots, y_m \epsilon L$. This contradicts the maximality of the independent set $\{u_1, \cdots, u_m\}$. Hence $k \leqslant m$ and statement (i) is proved.

Statement (ii) follows immediately from (i).

In order to prove Theorem 2 below we need the following lemma whose proof is quite simple and hence is ommitted.

Lemma 1. *Let $A \epsilon \mathfrak{M}(n, F)$. Then rank $Z_F(A) =$ rank $Z_F(UAU^*)$ for all nonsingular U in $\mathfrak{M}(n, F)$.*

Theorem 2. *Let $S \epsilon \mathfrak{M}(n, F)$ be skew-hermitian.*

(i) *If $F = R$, then rank $Z_F(S) = n - \frac{1}{2}$ rank S.*

(ii) *If $F = C$, then rank $Z_F(S) = n -$ rank $S + \min\{p, q\}$, where p, q are respectively the numbers of positive and negative eigenvalues of the hermitan matrix $\varepsilon_1 S$, where $\varepsilon_1 = \sqrt{-1}$ (ϵC).*

(iii) *If $F = H$, then*

$$\text{rank } Z_F(S) = \begin{cases} n - \frac{1}{2} \text{ rank } S, & \text{if rank } S \text{ is even}, \\ n - \frac{1}{2}(1 + \text{rank } S), & \text{if rank } S \text{ is odd}. \end{cases}$$

Proof. *Case* 1. $F = R$.

By a well-known result (for example, see [3, p. 285]) and Lemma 1, we may assume that

$$S = \text{diag}\{0_k, J_1, \cdots, J_l\},$$

where 0_k is the $k\times k$ zero matrix and each $J_i=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$. Then rank $S=2l$ and $n=k+2l$. Let

$$u_j=(0,\cdots,0,\underset{\uparrow\ j\text{-th component}}{1},0,\cdots,0),\qquad j=1,\cdots,k,$$

and

$$u_{k+i}=(0,\cdots,0,\underset{\uparrow\ (k+2i-1)\text{ th component}}{1},1,0,\cdots,0),\ i=1,\cdots,l.$$

Then $\{u_1,\cdots,u_{k+l}\}$ is an independent set in $z_F(S)(=R^n)$ and $u_s Su_t^*=0$ for all $s,t=1,\cdots,k+l$. Let $u=(r_1,\cdots,r_n)\in R^n$ be such that $uSu_s^*=0$ for all $s=1,\cdots,k+l$. Then

$$-r_{k+2i-1}+r_{k+2i}=0,\ i=1,\cdots,l.$$

Hence u is linearly dependent on $\{u_1,\cdots,u_{k+l}\}$, and by Theorem 1, rank $Z_F(S)=k+l=n-l=n-\frac{1}{2}$ rank S.

Case 2. $F=C$.

Let p,q and ε_1 be defined as in statement (ii). Then by a well-known result (for example, see [3, p. 274]) and Lemma 1, we may assume that

$$S=\mathrm{diag}\ \{0_k,-\varepsilon_1 I_p,\varepsilon_1 I_q\},$$

where 0_k is the $k\times k$ zero matrix and I_p and I_q are respectively the $p\times p$ and $q\times q$ identity matrices. Then rank $S=p+q$ and $n=k+p+q$. Suppose $p\leqslant q$. Let

$$u_j=(0,\cdots,0,\underset{\uparrow\ j\text{-th component}}{1},0,\cdots,0), j=1,\cdots,k,$$

and

$$u_{k+i}=(0,\cdots,0,\underset{\uparrow\ (k+i)\text{ th component}}{1},0,\cdots,0,\underset{\uparrow\ (k+p+i)\text{ th component}}{1},0,\cdots,0), i=1,\cdots,p.$$

Then $\{u_1,\cdots,u_{k+p}\}$ is an independent set in $z_F(S)$ and $u_s Su_t^*=0$ for all $s,t=1,\cdots,k+p$. Let $u=(c_1,\cdots,c_n)\in C^n$ such that

(1) $$uSu_s^*=0, \text{ for all } s=1,\cdots,k+p,$$

and

(2) $$uSu^*=0.$$

From (1) it follows that

(3) $$-c_{k+i}\,\varepsilon_1+c_{k+p+i}\,\varepsilon_1=0,\qquad i=1,\cdots,p,$$

and from (2) and (3) it follows that

$$\sum_{l=1}^{q-p} c_{k+2p+l}\, \varepsilon_1\, c^*_{k+2p+l}=0.$$

Hence we have $c_{k+i}=c_{k+p+i}$ for all $i=1,\cdots,p$ and $c_{k+2p+l}=0$ for all $l=1,\cdots,q-p$, and consequently u is linearly dependent on $\{u_1,\cdots,u_{k+p}\}$. By Theorem 1 we see that rank $Z_F(S)=k+p=n-$rank $S+\min\{p,q\}$.

Case 3. $F=H$,

Let rank $S=m$ and $k=n-m$. Then by a known result (for example, see [4] or [5]) and Lemma 1, we may assume that

$$S=\begin{cases}\operatorname{diag}\{0_k, -\varepsilon_1 I_{m/2}, \varepsilon_1 I_{m/2}\}, & \text{if } m \text{ is even,}\\ \operatorname{diag}\{0_k, -\varepsilon_1 I_{(m-1)/2}, \varepsilon_1 I_{(m-1)/2}, \varepsilon_1\}, & \text{if } m \text{ is odd,}\end{cases}$$

and by proceeding as in Case 2, we can easily prove statement (iii).

3. Some results of V. R. Tihomirov.

Let $A=(a_{ij})\in\mathfrak{M}(n,F)$. As in Tihomirov's papers [1, 2], we define $h(A)=\sum_{i=1}^{n}\sum_{j=1}^{n}[a_{ij}-a^*_{ji}][a^*_{ij}-a_{ji}]$. Obviously, we have $h(A)=Tr[(A-A^*)(A^*-A)]$, where Tr means "*trace*". In order to prove Theorems 3 and 4 we need the following lemma.

Lemma 2. *Let* $X, Y\in\mathfrak{M}(n,F)$. *Then*

(i) $Tr(XX^*)\geqslant 0$ *and* $=0$ *if and only if* $X=0$;

(ii) $Re\ Tr(XY)=Re\ Tr(YX)$, *where Re means "the real part", and*

(iii) $XYX^*=0$ *if and only if* $Tr(X^*XYX^*XY^*)=0$.

Proof. (i) and (ii) follow immediately from definition. We now prove (iii). If $XYX^*=0$, then $X^*XYX^*XY^*=0$, so that $Tr(X^*XYX^*XY^*)=0$. Conversely, if $Tr(X^*XYX^*XY^*)=0$, then, by (ii), $Re\ Tr(XYX^*XY^*X^*)=Re\ Tr(X^*XYX^*XY^*)=0$, and consequently by (i) $XYX^*=0$.

Theorem 3. *Let* $A\in\mathfrak{M}(n,F)$. *Then*

$$h(A)=h(UAU^*),$$

for any unitary matrix U *in* $\mathfrak{M}(n,F)$.

Proof. By Lemma 2 and the fact that U is unitary, we have

$$\begin{aligned}h(UAU^*)&=Re\ Tr[(UAU^*-UA^*U^*)(UA^*U^*-UAU^*)]\\&=Re\ Tr[U(A-A^*)(A^*-A)U^*]\\&=Re\ Tr[U^*U(A-A^*)(A^*-A)]\\&=Re\ Tr[(A-A^*)(A^*-A)]\\&=h(A).\end{aligned}$$

Theorem 4. *Let* $A, X\in\mathfrak{M}(n,F)$, $X^*X=I_n+E$ *and* $S=A-A^*$.

(i) *If XAX^* is hermitian, then*

$$h(A)=Tr(ESS)+Tr(SES)+Tr(ESES). \tag{4}$$

(ii) If

$$h(A)=2\ Re\ Tr(ESS)+Re\ Tr(ESES), \tag{5}$$

then XAX^* is hermitian.

Proof. If XAX^* is hermitian, then $XSX^*=0$ and by Lemma 2 and the fact that S is skewhermitian, we have $Tr(X^*XSX^*XS)=0$. Hence

$$-h(A)+Tr(ESS)+Tr(SES)+Tr(ESES)=0.$$

If (5) holds, then by Lemma 2 we have

$$Re\ Tr(SS)+Re\ Tr(ESS)+Re\ Tr(SES)+Re\ Tr(ESES)=0.$$

Hence

$$Re\ Tr(X^*XSX^*XS)=0,$$

and by Lemma 2 we have $XSX^*=0$ and hence XAX^* is hermitian.

Remarks.

1. Since for $F=R$ or C, we have $Tr(XY)=Tr(YX)$ for any $X, Y\in \mathfrak{M}(n,F)$ and $Tr(XY)$ is real if X and Y are hermitian, and since E and SS are, by definition, hermitian, so in these cases we see that $Tr(ESS)(=Tr(SES))$ and $Tr(ESES)$ are real in expression (4). But this is not the case for $F=H$. For example, take $A=\begin{pmatrix}0 & \varepsilon_1+\varepsilon_2\\ 0 & -\varepsilon_3\end{pmatrix}$ and $X=\begin{pmatrix}\varepsilon_1 & 1\\ 0 & 0\end{pmatrix}$, where $\{1, \varepsilon_1, \varepsilon_2, \varepsilon_3\}$ is the basis of H, then $S=A-A^*=\begin{pmatrix}0 & \varepsilon_1+\varepsilon_2\\ \varepsilon_1+\varepsilon_2 & -2\varepsilon_3\end{pmatrix}$, $E=X^*X-I_2=\begin{pmatrix}0 & -\varepsilon_1\\ \varepsilon_1 & 0\end{pmatrix}$, $XAX^*=\begin{pmatrix}-1 & 0\\ 0 & 0\end{pmatrix}$(hermitian), $h(A)=8$, $Tr(ESS)=4\varepsilon_3+4$, $Tr(SES)=4$ and $Tr(ESES)=-4\varepsilon_3$.

2. The mistake in *Tihomirov's* paper[2] is that, besides that there are misprints in indices, the term $Tr(ESS)+Tr(SES)$ has been omitted in expressions (4) and (5).

REFERENCES

[1] V. R. Tihomirov, ***An invariant of a unitary transformation*** (Russian), Kabardino–Balkarsk. Gos. Univ. Ucen. Zap. 24 (1965), 275–277.

[2] V. R. Tihomirov, ***A necessary and sufficientc ondition for the conversion of an arbitrary matrix into a Hermitian matrix as a result of a similar transformation*** (Russian), Kabardino-Balkarsk. Gos. Univ. Ucen. Zap. 30 (1966), 263–264.

[3] F. R. Gantmacher, ***The Theory of Matrices***, vol. 1, Chelsea, 1959.

[4] J. Radon, ***Lineare Scharen orthogonaler Matrizen***, Abh. Math. Sem. Univ. Hamburg. 1 (1922), 1–14.

[5] H. C. Lee, ***Eigenvalues and canonical forms of matrices with quaternion coefficients***, Proc. Roy. Irish Acad. Sect. A 52 (1949), 253–260.

UNIVERSITY OF HONG KONG, HONG KONG.