Kyungpook Math. J.
Volume 21, Number 1.
June, 1981.

SOME REMARKS ON UNIPOTENT MATRICES

By Chong-Yun Chao

Putname and Wintner in [2] gave a condition under which A commuting with $AB-BA$ implies the spectrum of $ABA^{-1}B^{-1}$ is only 1 where A and B are bounded, linear and nonsingular operators of a Hilbert space. In [1], Herstein proved the following theorem:

Let A and B be a regular $n\times n$ matrices over a field F of characteristic a prime $p>n$. If

$$A(AB-BA)=(AB-BA)A, \tag{1}$$

then $AB^{-1}A^{-1}B$ is unipotent (i.e., $AB^{-1}A^{-1}B-I$ is nilpotent where I is the identity matrix).

Apparently, Herstein's theorem also holds for the case F being of characteristic zero.

Here in §1, we reestablish Herstein's theorem. Although the method is similar, it seems that our computation is simpler. In §2, we use Herstein's theorem and I. Schur's work, namely the well known Schur's lemma and the centralizers of a permutations group, to obtain some theorems concerning matrix commutators and centralizers.

1. Unipotent Matrix

THEOREM 1. *Let A and B be regular $n\times n$ matrices over a field F whose characteristic is either zero or a prime $p>n$. If the condition* (1) *holds, then $A^tBA^{-t}B^{-1}$ is unipotent for any integer t.*

PROOF. We shall consider the case of $t=1$. Let $U=ABA^{-1}B^{-1}$. Then since similar matrices have the same eigenvalues, U and AUA^{-1} have the same eigenvalues. By using the condition (1),

$AUA^{-1}=A(ABA^{-1}B^{-1})A^{-1}=(A^2B)A^{-1}B^{-1}A^{-1}=(2ABA-BA^2)A^{-1}B^{-1}A^{-1}\ 2I-BAB^{-1}A^{-1}=2I-U^{-1}$.

That is, U and $2I-U^{-1}$ have the same eigenvalues. Let $\mu_1, \mu_2, \cdots, \mu_n$ be the eigenvalues of U in the algebraic closure of F. Then the eigenvalues of $2I-U^{-1}$ are $2-\mu_i^{-1}$ for $i=1, 2, \cdots, n$, and we have, for $1\leq k\leq n$,

$$\mu_1=2-\mu_2^{-1},\ \mu_2=2-\mu_3^{-1},\cdots,\ \mu_{k-1}=2-\mu_k^{-1},\ U_k=2-\mu_1^{-1}$$

where the subscripts of μ's, if necessary, may be rearranged. Consequently, by substitutions, we have

$$\mu_1=\frac{k\mu_k-(k-1)}{(k-1)\mu_k-(k-2)} \tag{2}$$

holds since the characteristic is either zero or $p>n\geq k$. Again, substituting $\mu_k=2-\mu_1^{-1}$ into (2), we have

$$\mu_1=\frac{(k+1)\mu_1-k}{k\mu_1-(k-1)}. \tag{3}$$

After simplifying (3), we obtain $k(\mu_1-1)^2=0$. Hence, $\mu_1=1$. Conseqently, $\mu_2=\mu_3=\cdots\mu_k=1$. If $k<n$, we repeat the same process for a finite number of times to obtain $\mu_1=\mu_2=\cdots=\mu_n=1$.

For the case of t being a positive integer, we claim that

$$A^tBA^{-t}B^{-1}=tU-(t-1)\ I. \tag{4}$$

By induction. Clearly, (4) holds for $t=1$. Assume

$$A^{t-1}BA^{-(t-1)}B^{-1}=(t-1)U-(t-2)I. \tag{5}$$

After multiplying A^{t-2} *to* (1), we have

$$A^{t-1}(AB-BA)=A^{t-2}(AB-BA)A=(AB-BA)A^{t-1},$$

and

$$A^t B=A^{t-1}BA+ABA^{t-1}-BA^t. \tag{6}$$

By using (6) and (5), we have

$$\begin{aligned}A^tBA^{-t}B^{-1}&=(A^{t-1}\ BA+ABA^{t-1}-BA^t)A^{-t}B^{-1}\\&=A^{t-1}BA^{-(t-1)}B^{-1}+ABA^{-1}B^{-1}-I\\&=(t-1)U-(t-2)I+U-I=tU-(t-1\ )I.\end{aligned}$$

Since the eigenvalues of U are 1, it follows that all the eigenvalues of $A^tBA^{-t}B^{-1}$ are $t-(t-1)=1$ for any positive integer t.

For the case of $t=-s$ being a negative integer, we claim that

$$A^{-s}BA^sB^{-1}=(s+1)I-sU. \tag{7}$$

By induction. For $s=1$, let $W=A^{-1}BAB^{-1}$. Since (1) is also

$$(AB-BA)A^{-1}=A^{-1}(AB-BA),\ A^{-1}BA=2B-ABA^{-1}$$

and $W=A^{-1}BAB^{-1}=(2B-ABA^{-1})B^{-1}=2I-U$. Assume

$$A^{-s+1}BA^{s-1}B^{-1}=sI-(s-1)U. \tag{8}$$

After multiplying A^{-s+1} to (1) in the form $(AB-BA)A^{-1}=A^{-1}(AB-BA)$, we have

$$(AB-BA)A^{-s}=A^{-s}(AB-BA),$$

and

$$A^{-s}BA=A^{-s+1}B+BA^{-s+1}-ABA^{-s} \tag{9}$$

By using (9) and (8), we have

$$\begin{aligned} A^{-s}BA^sB^{-1} &= (A^{-s}BA)\,A^{s-1}B^{-1} \\ &= (A^{-s+1}B+BA^{-s+1}-ABA^{-s})A^{s-1}B^{-1} \\ &= A^{-s+1}BA^{s-1}B^{-1}+I-ABA^{-1}B^{-1} \\ &= sI-(s-1)U+I-U=(s+1)I-sU. \end{aligned}$$

Again, since the eigenvalues of U are 1, it follows that all the eigenvalues of $A^{-s}BA^sB^{-1}$ are $(s+1)-s=1$ for any positive integer s, i.e., all the eigenvalues of $A^tBA^{-t}B^{-1}$ are 1 for any negative integer t.

COROLLARY 1.1 *Let A and B be the same as in Theorem 1. If A commutes with $AB-BA$, then* (a) $BA^{-t}B^{-1}A^t$, $A^{-t}B^{-1}A^tB$ and $B^{-1}A^tBA^{-t}$ *are unipotent for any integer t.* (b) $(A^tBA^{-t}B^{-1})^{t'}$, $(BA^{-t}B^{-1}A^t)^{t'}$, $(A^{-t}B^{-1}A^tB)^{t'}$ and $(B^{-1}A^tBA^{-t})^{t'}$ *are unipotent for any integer t and t'.*

PROOF. (a) Each of these matrices is similar to $A^tBA^{-t}B^{-1}$, i.e.,

$$\begin{aligned} &BA^{-t}B^{-1}A^t=(A^t)^{-1}(A^tBA^{-t}B^{-1})A^t, \\ &A^{-t}B^{-1}A^tB=(A^tB)^{-1}(A^tBA^{-t}B^{-1})(A^tB), \text{ and} \\ &B^{-1}A^tBA^{-t}=(A^tBA^{-t})^{-1}(A^tBA^{-t}B^{-1})(A^tBA^{-t}). \end{aligned}$$

Hence, it follows from Theorem 1 that each of them is unipotent.

(b) Using the facts that the product of two unipotent matrices is again unipotent if they commute, and that the inverse of a unipotent matrix is unipotent, (b) follows.

THEOREM 2. *Let A and B be the same as in Theorem 1. If B commutes with $AB-BA$, then $AB^tA^{-1}B^{-t}$ is unipotent for any integer t.*

PROOF. Since B commutes with $AB-BA$, B commutes with $BA-AB$. By Theorem 1, $B^tAB^{-t}A^{-1}$ is unipotent for any integer t. Since the inverse of a unipotent matrix is unipotent, $(B^tAB^{-t}A^{-1})^{-1}=AB^tA^{-1}B^{-t}$ is unipotent for any integer t.

COROLLARY 2.1 *Let A and B be the same as in Theorem 1. If B commutes with $AB-BA$, then* (a) $B^tA^{-1}B^{-1}A$, $A^{-1}B^{-t}AB^t$ *and* $B^{-t}AB^tA^{-1}$ *are unipotent for any integer t.* (b) $(AB^tA^{-1}B^{-t})^{t'}$, $(B^tA^{-1}B^{-t}A)^{t'}$, $(A^{-1}B^{-t}AB^t)^{t'}$ *and* $(B^{-t}AB^tA^{-1})^{t'}$ *are unipotent for any integers t and t'.*

REMARKS. (1) Our Theorem 1 can also be proved by a similar method to the Herstein's proof in [1]. But our proof above seems to be simpler. (2) The following example shows that, although A and B are regular $n\times n$ matrices over a field F of characteristic either zero or $p>n$ and A commutes with $AB-BA$, $AB^2A^{-1}B^{-2}$ is not necessarily unipotent:

Let $A=\begin{bmatrix}1&0&0\\0&0&1\\0&-1&2\end{bmatrix}$ and $B=\begin{bmatrix}1&1&0\\1&1&-1\\0&-1&1\end{bmatrix}$. Then $B^2=\begin{bmatrix}2&2&-1\\2&3&-2\\-1&-2&2\end{bmatrix}$.

$AB-BA=\begin{bmatrix}0&1&-1\\-1&-2&2\\-1&-2&2\end{bmatrix}$ and $A(AB-BA)=(AB-BA)A$ holds.

The eigenvalues of $ABA^{-1}B^{-1}=\begin{bmatrix}1&0&-1\\0&0&1\\0&-1&2\end{bmatrix}$ are 1, 1 and 1, and the eigenvalues of $AB^2A^{-1}B^{-2}=\begin{bmatrix}0&1&0\\0&0&1\\1&-2&2\end{bmatrix}$ are 1, $\frac{1}{2}+\frac{\sqrt{3}i}{2}$ and $\frac{1}{2}-\frac{\sqrt{3}i}{2}$.

(3) Similarly, B commuting with $AB-BA$ does not necessarily imply the unipotency of $A^2BA^{-2}B^{-1}$.

THEOREM 3. *Let A and B be the same as in Theorem 1. (1) If A commutes with $AB-BA$ and if $U=ABA^{-1}B^{-1}$ and $B^{-1}UB$ commute, then $AB^2A^{-1}B^{-2}$ is unipotent. (2) If B commutes with $AB-BA$ and if U and $A^{-1}UA$ commte, then $A^2BA^{-2}B^{-1}$ is unipotent.*

PROOF. (1) $B^{-1}(AB^2A^{-1}B^{-2})B=(B^{-1}ABA^{-1})(ABA^{-1}B^{-1})$
$=(B^{-1}(ABA^{-1}B^{-1})B)(ABA^{-1}B^{-1})=(B^{-1}UB)U.$

Since $AB^2A^{-1}B^{-2}$ and $B^{-1}(AB^2A^{-1}B^{-2})B$ have the same eigenvalues and since U is unipotent by Theorem 1 and since U and $B^{-1}UB$ commute, $AB^2A^{-1}B^{-2}$ is unipotent.

(2) $A^{-1}(A^2BA^{-2}B^{-1})A=(ABA^{-1}B^{-1})(BA^{-1}B^{-1}A)$

$$=(ABA^{-1}B^{-1})(A^{-1}(ABA^{-1}B^{-1})A)=U(A^{-1}UA).$$

Since $A^2BA^{-2}B^{-1}$ and $A^{-1}(A^2BA^{-2}B^{-1})A$ have the same eigenvalues and since U is unipotent by Theorem 2 and since U and $A^{-1}UA$ commute, $A^2BA^{-2}B^{-1}$ is unipotent.

2. The centralizers of a matrix

Let $M_n(F)$ be the algebra of the $n\times n$ matrices over F of characteristic either zero or a prime $p>n$, A be a matrix in $M_n(F)$, $C(A)=\{D\in M_n(F) : AD=DA\}$ and $V(A)=\{E\in M_n(F) : A(AE-EA)=(AE-EA)A\}$. Then $C(A)$ is a subalgebra of $M_n(F)$, $V(A)$ is a subspace of $M_n(F)$ such that $C(A)\subseteq V(A)$. To find a necessary and sufficient condition for $C(A)=V(A)$ seems to very difficult. Here, we shall use I. Schur's work, namely, the well known Schur's lemma and his result in [4], to prove the following theorems.

THEOREM 4. *Let A be a regular matrix in $M_n(F)$ such that the characteristic polynomial of A is irreducible over F. Then the regular matrices together with the zero matrix in $V(A)$ constitute a field. In fact, the field is $C(A)$.*

PROOF. Let B be an arbitrary regular matrix in $V(A)$. Then B satisfies the condition (1) and, by Theorem 1, $ABA^{-1}B^{-1}$ is unipotent, i.e., $ABA^{-1}B^{-1}=I+N$ where N is a nilpotent matrix. In fact, we have

$$(AB-BA)A^{-1}B^{-1}=N$$

which implies that $AB-BA$ must be a singular matrix in $C(A)$. By Theorem 8 in [5], which states that $C(A)$ is a field if and only if the characteristic polynomial is irreducible over F, $AB-BA$ must be the zero matrix, i.e., $AB=BA$ and B belongs to $C(A)$. Since $C(A)\subseteq V(A)$, the regular matrices in $V(A)$ together with the zero matrix constitute the field $C(A)$.

THEOREM 5. *Let $\{A_1, A_2, \cdots, A_k\}$ be an irreducible set of regular matrices in $M_n(F)$. Then the regular matrices together with the zero matrix in $\bigcap_{i=1}^{k} V(A_i)$ constitute a division ring. In fact, this division ring is $\bigcap_{i=1}^{k} C(A_i)$.*

PROOF. Let B be an arbitrary regular matrix in $\bigcap_{i=1}^{k} V(A_i)$. Then $A_i(A_iB-BA_i)=(A_iB-BA_i)A_i$ for $i=1,2,\cdots,k$, and by Theorem 1, $A_iBA_i^{-1}B^{-1}$ is unipotent, i.e., $A_iBA_i^{-1}B^{-1}=I+N_i$ where N_i is a nilpotent matrix. Hence

$$(A_iB-BA_i)A_i^{-1}B^{-1}=N_i$$

for $i=1, 2, \cdots, n$ imply that A_iB-BA_i must be a singular matrix in $\bigcap_{i=1}^{k} C(A_i)$. Since $\bigcap_{i=1}^{k} C(A_i)$ is a division ring by Schur's lemma, A_iB-BA_i must be a zero matrix, i.e., $A_iB=BA_i$ for $i=1,2,\cdots,k$ and $B\in \bigcap_{i=1}^{k} C(A_i)$. Since $\bigcap_{i=1}^{k} C(A_i)\subseteq \bigcap_{i=1}^{k} V(A_i)$, the regular matrices in $\bigcap_{i=1}^{k} V(A_i)$ together with the zero matrix constitute the division ring $\bigcap_{i=1}^{k} C(A_i)$.

THEOREM 6. *Let σ be a permutation on n letters and P be its corresponding permutation matrix $M_n(F)$ where the characteristic of F is zero. Then $V(P)=C(P)$.*

PROOF. By using the Theorem 1 in [4], we know that if $C=(c_{ij})\in C(P)$. Then, for $i, j=1, 2, \cdots, n$,

$$c_{ij}=c_{\sigma i,\sigma j}=c_{\sigma^2 i,\sigma^2 j}=\cdots=c_{\sigma^k i,\sigma^k j}$$

where the order of σ is $k+1$.

Let $P=(p_{ij})$ and $B=(b_{ij})$ be an arbitrary matrix in $V(P)$. Then, by using the properties of a permutation matrix, we have, for $i,j=1,2,\cdots,n$,

$$\begin{aligned}(PB-BP)_{ij}&=\sum_{k=1}^{n} p_{ik}\,b_{kj}-\sum_{k=1}^{n} b_{ik}p_{kj}\\ &=p_{i,\sigma i}\,b_{\sigma i,j}-b_{i,\sigma^{-1}j}\,p_{\sigma^{-1}j,j}\\ &=b_{\sigma i,j}-b_{i,\sigma^{-1}j}. \qquad (10)\end{aligned}$$

Since $(PB-BP)\in C(P)$, $(PB-BP)_{ij}=(PB-BP)_{\sigma i,\sigma j}=(PB-BP)_{\sigma^2 i,\sigma^2 j}=\cdots=(PB-BP)_{\sigma^k i,\sigma^k j}$. Then by (10), we have

$$\begin{aligned}&b_{\sigma i,j}-b_{i,\sigma^{-1}j}=b_{\sigma^2 i,\sigma j}-b_{\sigma i,j},\\ &b_{\sigma i,\sigma j}-b_{\sigma i,j}=b_{\sigma^3 i,\sigma^2 j}-b_{\sigma^2 i,\sigma j},\\ &b_{\sigma^3 i,\sigma^2 j}-b_{\sigma^2 i,\sigma j}=b_{\sigma^4 i,\sigma^3 j}-b_{\sigma^3 i,\sigma^2 j},\\ &\quad\vdots\qquad\qquad\vdots\qquad\qquad\vdots\\ &b_{\sigma^{k+1} i,\sigma^k j}-b_{\sigma^k i,\sigma^{k-1} j}=b_{\sigma i,j}-b_{i,\sigma^{-1}j}.\end{aligned}$$

That is,

$$\begin{aligned}&2b_{\sigma i,j}=b_{\sigma^2 i,\sigma j}+b_{i,\sigma^k j},\\ &2b_{\sigma^2 i,\sigma j}=b_{\sigma^3 i,\sigma^2 j}+b_{\sigma i,j},\end{aligned}$$

$$2b_{\sigma^3 i,\sigma^2 j}=b_{\sigma^4 i,\sigma^3 j}+b_{\sigma^2 i,\sigma j}$$
$$\vdots \qquad \vdots$$
$$2b_{i,\sigma^k j}=b_{\sigma i,\, j}+b_{\sigma^k i,\sigma^{k-1} j}$$

Solving these equation simultaneously, we have

$b_{\sigma i,\, j}=b_{\sigma^2 i,\sigma j}=b_{\sigma^3 i,\sigma j^2}=\cdots=b_{i,\sigma^k j}$ for all i, $j=1, 2, \cdots, n$, i.e., $B\in C(P)$ *and* $V(P)=C(P)$.

The following example shows that if the characteristic of F is not zero, our Theorem 6 does not hold: Let F be the field of 3 elements and

$P=\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$. Then $C(P)=\left\{\begin{bmatrix} a & b & c \\ c & a & b \\ b & c & a \end{bmatrix} : a, b, c\in F\right\}$. Let $B=\begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$. Then $PB-BP=\begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $P(PB-BP)=(PB-BP)P$ hold. Hence, $B\in V(P)$, but $B\notin C(P)$, i.e., $C(P)\subsetneq V(P)$.

Department of Mathematics,
University of Pittsburgh,
Pittsburgh, PA 15260

REFERENCES

[1] I. N. Herstein, *On a theorem of Putnam and Wintner*, Proc. Amer. Math. Soc. Vol. 9(1958), 363—364.

[2] C. Putnam and A. Wintner, *On the spectra of commutators*, Proc. Amer. Math. Soc. Vol. 9(1958), 360—362.

[3] D. W. Robinson, *A note on matrix commutators*, Canad. J. Math. 17 (1965), 527—532.

[4] I. Schur, *Zur Theorie der einfach transitiven Permutations gruppen*, S. B. Preuss. Akad. Wiss. Phys. -Math. Kl (1933), 598—623.

[5] D. A. Suprenenko and R. I. Tyshkevich, *Commutative Matrices*, Academic Press, 1968.