On idempotent matrices.

By NISAR A. KHAN in Aligarh.

1. A square matrix $A=(a_{ij})$ is called idempotent if $A^2=A$ (see [1], p. 88.). In [2] SHAH and ANSARI have proved certain results for nilpotent matrices. The purpose of this paper is to establish analogous results for idempotent matrices. Let k be an integer >1; we shall denote by A an idempotent (square) matrix of order $n>1$, with elements a_{ij} belonging to the field F of real numbers, a_{ij} being of the form $P/2^Q$ where P and Q are integers (positive, negative or zero) such that $P\equiv 0 \pmod{k}$ and $Q\leqq k$.

If X and Y be two matrices such that the elements of $X-Y$ are (integer) multiples of k, then we write $X\equiv Y \pmod{k}$. Also I denotes an identity matrix, and 0 a null matrix whose orders will be clear from the context. Following LEHMER we shall call an integer p a pseudoprime if $2^p\equiv 2 \pmod{p}$ and p is not a prime. (For results on pseudoprimes see [3] and references given therein.)

By simple calculations it can be shown that

(i) If A is idempotent, so is the matrix $(I-A)^k$, k being a positive integer,

(ii) the idempotent matrices A and $(I-A)^k$ are orthogonal, and

(iii) if the rank of A is r, the rank of $(I-A)^k$ is $n-r$.

2. We now prove the following

Theorem 1. *If k is a prime or a pseudoprime, then*

$$\{2(I+A)\}^k\equiv 2I \pmod{k},$$

and conversely.

PROOF. If k is a prime or a pseudoprime, then $2^k\equiv 2 \pmod{k}$; and by hypothesis the elements of the matrix 2^kA are all integer multiples of k. Hence

$$\begin{aligned}\{2(I+A)\}^k-2I &= 2^k(I+{}^kc_1A+{}^kc_2A^2+\cdots+{}^kc_kA^k)-2I=\\ &=(2^k-2)I+2^k({}^kc_1+{}^kc_2+\cdots+{}^kc_k)A=\\ &=(2^k-2)I+2^k(2^k-1)A\equiv\\ &\equiv 0 \pmod{k}.\end{aligned}$$

Conversely, the elements of the matrix $\{2(I+A)\}^k-2I$ or $(2^k-2)I+ +2^k(2^k-1)A$ are multiples of k. Therefore,

$$2^k-2\equiv 0 \pmod k$$

or

$$2^k\equiv 2 \pmod{k}.$$

Hence k is either a prime or it must be a pseudoprime.

By using another important property of the binomial coefficients we can similarly prove the following

Theorem 2. *If k is a prime or a pseudoprime, then*

$$\{2(I-A)\}^k\equiv 2I \pmod{\mathrm{k}},$$

and conversely.

3. We now establish the following theorems regarding the characteristic roots of the polynomials $f(I\pm A)$.

Theorem 3. *If $f(x)$ is any polynomial in x with scalar coefficients belonging to F, then the characteristic roots of $f(I+A)$ are $f(2)$ with multiplicity r and $f(1)$ with multiplicity $n-r$, where r is the rank of A.*

The proof of the above theorem will be facilitated by the following lemma:

Lemma. *For an idempotent matrix A the characteristic polynomial is $|A-\lambda I|=(1-\lambda)^r(-\lambda)^{n-r}$, where r is the rank of A.*

Proof. It is well known that an idempotent matrix A is similar to $A_r=I_r\dotplus 0_{n-r}$, where r is the rank of A and $\dotplus$ denotes the direct sum. Further, two similar matrices have the same characteristic equation and hence the same characteristic polynomial. Therefore,

$$|A-\lambda I|=|A_r-\lambda I|=(1-\lambda)^r(-\lambda)^{n-r},$$

r being the rank of A.

This proves the lemma.

It may be remarked here that the converse is not true. This is illustrated by the following matrix:

$$A=\begin{pmatrix}1&0&2\\0&0&0\\0&0&1\end{pmatrix}, \text{ so that } A^2=\begin{pmatrix}1&0&4\\0&0&0\\0&0&1\end{pmatrix}\neq A,$$

i. e., A is not idempotent, but $|A-\lambda I|=(-\lambda)(1-\lambda)^2$.

PROOF OF THEOREM 3. Let $f(x)=a_0+a_1x+a_2x^2+\cdots+a_Nx^N$.

$$\begin{aligned} f(I+A) &= a_0I+a_1(I+A)+a_2(I+A)^2+\cdots+a_N(I+A)^N= \\ &= (a_0+a_1+\cdots+a_N)I+(2-1)a_1A+(2^2-1)a_2A+\cdots+(2^N-1)a_NA= \\ &= I\cdot f(1)+A\{f(2)-f(1)\} \end{aligned}$$

$$\begin{aligned} f(I+A)-\lambda I &= \{f(1)-\lambda\}I+\{f(2)-f(1)\}A \\ &= \{f(2)-f(1)\}A-\{\lambda-f(1)\}I \\ &= lA-mI, \text{ where } l=f(2)-f(1), m=\lambda-f(1), \\ &= l\left(A-\frac{m}{l}I\right), \text{ if } l\neq 0. \end{aligned}$$

Now, since A is idempotent,

$$|f(I+A)-\lambda I|=\left|l\left(A-\frac{m}{l}I\right)\right|=l^n\left|A-\frac{m}{l}I\right|=l^n\left(1-\frac{m}{l}\right)^r\left(-\frac{m}{l}\right)^{n-r}.$$

Thus the characteristic roots of $f(I+A)$ are given by

$$l^n\left(1-\frac{m}{l}\right)^r\left(-\frac{m}{l}\right)^{n-r}=0.$$

Since $l\neq 0$, *either* $\left(1-\frac{m}{l}\right)^r=0$, whence $m=l$, r times; i. e., $f(1)-f(2)=f(1)-\lambda$, r times, $\lambda=f(2)$ with multiplicity r; *or* $m^{n-r}=0$, whence $\lambda=f(1)$ with multiplicity $n-r$.

If $l=0$, or $f(1)=f(2)$, $|f(I+A)-\lambda I|=|\{\lambda-f(1)\}I|$, so that $\lambda=f(1)=f(2)$ with multiplicity n.

This completes the proof of the theorem.

In particular, the characteristic roots of $(I+A)^k$ are 2^k with multiplicity r and 1 with multiplicity $n-r$.

As an illustration we take the following numerical

Example. Let $A=\begin{pmatrix}1&0&0\\2&0&0\\3&0&0\end{pmatrix}$ be an idempotent matrix of rank 1. Then

$$(I+A)^5=I+31A=\begin{pmatrix}32&0&0\\62&1&0\\93&0&1\end{pmatrix}.$$

The characteristic roots of the above matrix are 32, 1 and 1.

Theorem 4. *If $f(x)$ is any polynomial in x with scalar coefficients belonging to F, then the characteristic roots of $f(I-A)$ are $f(0)$ with multiplicity r and $f(1)$ with multiplicity $n-r$, where r is the rank of A.*

PROOF. Let $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_N x^N$.

$$\begin{aligned} f(I-A) &= a_0 I + a_1(I-A) + a_2(I-A)^2 + \cdots + a_N(I-A)^N = \\ &= I \cdot f(1) - A\{f(1) - f(0)\}, \end{aligned}$$

$$\begin{aligned} |f(I-A) - \lambda I| &= |A\{f(0) - f(1)\} - I \cdot \{\lambda - f(1)\}| \\ &= |lA - mI|, \text{ where } l = f(0) - f(1) \text{ and } m = \lambda - f(1). \\ &= l^n \left(1 - \frac{m}{l}\right)^r \left(-\frac{m}{l}\right)^{n-r}, \text{ if } l \neq 0. \end{aligned}$$

The characteristic roots of $f(I-A)$ are given by

$$\left(1 - \frac{m}{l}\right)^r \left(-\frac{m}{l}\right)^{n-r} = 0.$$

If $\left(1 - \frac{m}{l}\right)^r = 0$, $m = l$ and $\lambda = f(0)$ with multiplicity r.

If $\left(\frac{m}{l}\right)^{n-r} = 0$, $\lambda = f(1)$ with multiplicity $n - r$.

But, if $l = 0$, i. e., $f(1) = f(0)$, the characteristic roots of $f(I-A)$ are given by $\{\lambda - f(1)\}^n = 0$, so that $\lambda = f(1) = f(0)$ with multiplicity n.

In particular, the characteristic roots of $(I-A)^k$ are 1 with multiplicity r and 0 with multiplicity $n - r$.

Theorem 5. *If $f(x)$ is a polynomial in x with coefficients belonging to F, and D is any nilpotent matrix of order n such that*

$$f(1)D + \{f(2) - f(1)\}DA = 0,$$

then the characteristic roots of $f(I+A) + D$ are $f(2)$ with multiplicity r and $f(1)$ with multiplicity $n - r$, where r is the rank of A.

PROOF. To prove this theorem we apply a theorem of REID [4]. The characteristic roots of $f(I+A) + D$ and $f(I+A)$ are the same if $D \cdot f(I+A) = 0$.

$$\begin{aligned} \text{But, } D \cdot f(I+A) &= D[a_0 I + a_1(I+A) + a_2(I+A)^2 + \cdots + a_N(I+A)^N] = \\ &= D[I \cdot f(1) + A\{f(2) - f(1)\}] = \\ &= f(1)D + \{f(2) - f(1)\}DA = \\ &= 0, \text{ by hypothesis.} \end{aligned}$$

Hence the characteristic roots of $f(I+A) + D$ are the same as those of $f(I+A)$. But, by Theorem 3, the characteristic roots of $f(I+A)$ are $f(2)$ and $f(1)$ with multiplicities r and $n - r$ respectively, where r is the rank of A.

This proves the theorem.

Corollary. *If D be any square nilpotent matrix of order n with elements in F such that $D+(2^k-1)DA=0$, then the characteristic roots of $(I+A)^k+D$ are 2^k and 1 with multiplicites r and $n-r$ respectively, where r is the rank of A.*

Theorem 6. *If D be any $n\times n$ nilpotent matrix with elements in F such that $D=DA$, then the characteristic roots of $(I-A)^k+D$ are 0 with multiplicity r and 1 with multiplicity $n-r$, where r is the rank of A.*

PROOF. With the help of the theorem by REID [4], we see that the characteristic polynomials of $(I-A)^k+D$ and $(I-A)^k$ are the same if $D(I-A)^k=0$.

But, $D(I-A)^k=D[I-{}^kc_1A+{}^kc_2A^2-\cdots+(-1)^k\,{}^kc_kA^k]=D(I-A)=$ $=D-DA=0$, by hypothesis.

Hence the characteristic roots of $(I-A)^k+D$ are the same as those of $(I-A)^k$.

But, by Theorem 4, the roots of $(I-A)^k$ are 0 with multiplicity r and 1 with multiplicity $n-r$.

This proves the theorem.

In general, let $g(x)=a_1x+a_2x^2+\cdots+a_Nx^N$ be a polynomial in x with coefficients belonging to F and with no term independent of x. Then

$$|g(I-A)+D-\lambda I|=|g(I-A)-\lambda I|,$$

for D is nilpotent and

$$\begin{aligned} D\cdot g(I-A)&=D[a_1(I-A)+a_2(I-A)^2+a_3(I-A)^3+\cdots+a_N(I-A)^N]=\\ &=D[a_1(I-A)+a_2(I-A)+a_3(I-A)+\cdots+a_N(I-A)]=\\ &=(a_1+a_2+\cdots+a_N)D(I-A)=\\ &=g(1)D(I-A)=0,\text{ by hypothesis.}\end{aligned}$$

Therefore, the characteristic roots of $g(I-A)+D$ are 0 with multiplicity r and $g(1)$ with multiplicity $n-r$, r being the rank of A.

Remark. To show that there exist matrices which satisfy the hypothesis of the above theorem, consider

$$A=\begin{pmatrix}1&0&0\\2&0&0\\3&6&0\end{pmatrix},$$ an idempotent matrix of order 3 and rank 1.

$$D=\begin{pmatrix}0&0&0\\0&0&0\\1&0&0\end{pmatrix},$$ a nilpotent matrix of order 3.

Here $D(I-A)=0$, also $(I-A)^5+D=I-A+D=\begin{pmatrix} 0 & 0 & 0 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$.

The characteristic roots of this matrix are 0, 1 and 1.

Finally, I wish to thank Professor S. M. SHAH for drawing my attention to this problem and for helpful criticism on this paper.

Bibliography.

[1] A. A. ALBERT, Modern Higher Algebra, *Chicago*, 1939.

[2] A. R. ANSARI and S. M. SHAH, A note on certain nilpotent matrices, *Math. Student* **20** (1952), 113—114.

[3] P. ERDŐS, On the converse of Fermat's Theorem, *Amer. Math. Monthly* **56** (1949) 623—624.

[4] W. T. REID, A note on the characteristic polynomials of certain matrices, *Proc. Amer. Math. Soc.* 1 (1950), 584—585.

(Received January 25, 1956.)