THE PROBABILITY THAT A MATRIX BE NILPOTENT

BY

N. J. FINE AND I. N. HERSTEIN[1]

In this paper we determine the number of nilpotent n by n matrices over (i) a finite field of characteristic p, and (ii) the integers modulo m. The results are most simple when expressed as probabilities by dividing by the total number of matrices in each case.

THEOREM 1. *The probability that an n by n matrix over $GF(p^\alpha)$ be nilpotent is $p^{-\alpha n}$.*

Proof. Let A be an n by n nilpotent matrix over the finite field F. Then[2] $V_n(F)$ has a basis $\{v_s^i\}$, $i = 1, \cdots, k$; $s = 1, \cdots, r_i$, such that

$$v_s^i A = v_{s-1}^i \qquad (1 \leqq i \leqq k; \quad 1 \leqq s \leqq r_i), \tag{1}$$

where it is understood that $v_0^i = 0$. Associated with each such A there is a partition π of n,

$$\pi : n = r_1 + r_2 + \cdots + r_k \qquad (r_1 \geqq r_2 \geqq \cdots \geqq r_k \geqq 1),$$

and two matrices are similar if and only if their corresponding partitions are identical. Let $g(\pi)$ be the number of matrices in the similarity class determined by π. Then the probability of nilpotence is

$$P = p^{-\alpha n^2} \sum_\pi g(\pi).$$

To determine $g(\pi)$, we select and fix a representative A of the similarity class belonging to π, together with a basis $\{v_s^i\}$ associated with A by (1). We then transform A by the ν nonsingular matrices over F to obtain all the elements of the class, each with multiplicity μ, where μ is the number of nonsingular matrices which commute with A. Then $g(\pi) = \nu/\mu$. Now it is known[3] that

$$\nu = x^{-n^2} f(n),$$

where $x = p^{-\alpha}$ and

$$f(n, x) = f(n) = (1 - x)(1 - x^2) \cdots (1 - x^n) \qquad (n \geqq 1)$$

$$f(0) = 1.$$

It remains to determine μ.

Received July 12, 1957.

[1] The first author wishes to acknowledge the support of the Air Force.

[2] See, for example, A. A. ALBERT, *Modern higher algebra*, University of Chicago Press, 1937, Chapter 4.

[3] L. E. DICKSON, *Linear groups*, Leipzig, 1901, p. 77.

Let B be an arbitrary matrix commuting with A. Then B is completely determined by its action on the vectors $\{v_{r_i}^i\}(1 \leqq i \leqq k)$. For if

$$v_{r_i}^i B = \sum_{j=1}^k \sum_{q=1}^{r_j} C_j^i(q) v_q^j \qquad (1 \leqq i \leqq k),$$

then for $s = 0, 1, 2, \cdots, r_i$,

$$v_{r_i-s}^i B = v_{r_i}^i A^s B = v_{r_i}^i B A^s = \sum_{j=1}^k \sum_{q=s+1}^{r_j} C_j^i(q) v_{q-s}^j .$$

In particular, for $s = r_i$, we find

$$0 = \sum_{j=1}^k \sum_{q=r_i+1}^{r_j} C_j^i(q) v_{q-r_i}^j ,$$

so $C_j^i(q) = 0$ for all q, i, j satisfying $r_i < q \leqq r_j$. In other words, we must have

$$\text{(2)} \qquad v_{r_i-s}^i B = \sum_{j=1}^k \sum_{q=s+1}^{m_{ij}} C_j^i(q) v_{q-s}^j \qquad (1 \leqq i \leqq k, \quad 0 \leqq s < r_i),$$

where $m_{ij} = \min(r_i, r_j)$. Conversely, given any set of constants

$$C_j^i(q), \qquad 1 \leqq i \leqq k, \qquad 1 \leqq j \leqq k, \qquad 1 \leqq q \leqq m_{ij},$$

the matrix B defined by (2) commutes with A. Therefore the number of such matrices is $p^{\alpha M}$, where

$$M = M(\pi) = \sum_{i,j=1}^k m_{ij} .$$

The parts r_i of the partition π can be grouped, so that the possible parts $n - u + 1$ $(u = 1, \cdots, n)$ appear with corresponding multiplicities b_u, which may be zero. With this convention, we may write

$$\pi : n = b_1 n + b_2(n-1) + \cdots + b_{n-1}\cdot 2 + b_n \cdot 1.$$

Then

$$M = \sum_{u,v=1}^n \sum_{\substack{r_i = n-u+1 \\ r_j = n-v+1}} \min(r_i, r_j) = \sum_{u,v=1}^n b_u\, b_v \min(n-u+1, n-v+1)$$

$$= \sum_{u=1}^n e_u(n-u+1),$$

where

$$e_u = b_u^2 + 2b_u \sum_{v=1}^{u-1} b_v = \left(\sum_{t=1}^u b_t\right)^2 - \left(\sum_{t=1}^{u-1} b_t\right)^2.$$

Thus, if we define

$$s_u = \sum_{t=1}^u b_t \qquad (u = 0, 1, 2, \cdots, n),$$

we have

$$\begin{aligned} M &= \sum_{u=1}^n (s_u^2 - s_{u-1}^2)(n-u+1) \\ &= \sum_{u=1}^n s_u^2(n-u+1) - \sum_{u=0}^n s_u^2(n-u), \\ M &= \sum_{u=1}^n s_u^2 . \end{aligned}$$

Of these $p^{\alpha M}$ matrices commuting with A, we must now find what proportion are nonsingular. We assert that if $AB = BA$, then B is nonsingular if and only if the vectors $\{v_1^iB\}$ are linearly independent. If B is nonsingular, the linear independence is obvious. Conversely, suppose that the $\{v_1^iB\}$ are linearly independent. Let $v \in V_n(F)$ be such that $vB = 0$, and write

$$v = \sum_{i=1}^k \sum_{q=1}^Q C_q^i v_q^i ,$$

with $C_Q^{i_0} \neq 0$ for some i_0. Applying A^{Q-1}, we find that

$$vA^{Q-1} = \sum_{i=1}^k C_Q^i v_1^i .$$

But

$$\sum_{i=1}^k C_Q^i(v_1^iB) = vA^{Q-1}B = vBA^{Q-1} = 0.$$

This contradicts the linear independence of $\{v_1^iB\}$ and proves our assertion.

If we put $s = r_i - 1$ in (2), we get

$$v_1^iB = \sum_{r_j \geqq r_i} C_j^i(r_i)v_1^j = \sum_{j \leqq i} C_j^i(r_i)v_1^j . \tag{3}$$

For $u = 1, 2, \cdots, n$, let V_u be the subspace spanned by those v_1^i for which $r_i = n - u + 1$. Thus V_u has dimension b_u, and if

$$W_u = V_1 \oplus V_2 \oplus \cdots \oplus V_u ,$$

then W_u has dimension $b_1 + b_2 + \cdots + b_u = s_u$. It is clear from (3) that $W_u B \subset W_u$, and that B is nonsingular if and only if

$$W_u B = W_u \qquad (u = 1, \cdots, n).$$

Let us define the linear transformation $\tilde{B}$ of W_n into itself by

$$v_1^i\tilde{B} = \sum_{r_j = r_i} C_j^i(r_i)v_1^j \qquad (i = 1, \cdots, k).$$

Clearly $V_u \tilde{B} \subset V_u$, and $\tilde{B}$ decomposes into a direct sum

$$\tilde{B}_1 \oplus \tilde{B}_2 \oplus \cdots \oplus \tilde{B}_n ,$$

where $\tilde{B}_u$ is defined on V_u by

$$v_1^i\tilde{B}_u = \sum_{r_j = n-u+1} C_j^i(r_i)v_1^j \qquad (r_i = n - u + 1).$$

Our next assertion is that B is nonsingular if and only if $\tilde{B}$ is also. To see this, let $wB = 0$, $w \neq 0$, $w \in W_n$, and write $w = w' + w''$, where

$$w' \in V_{u+1} , \qquad w'' \in W_u , \qquad w' \neq 0.$$

Then $w'B = -w''B \in W_u$, so $w'\tilde{B} = 0$ and $\tilde{B}$ is singular. Conversely, suppose that $w\tilde{B} = 0$, $w \neq 0$, $w \in W_n$. Making the same decomposition of w, we find that $w'\tilde{B} = -w''\tilde{B} \in W_u$, so $w'\tilde{B} = 0$. Hence $w'B \in W_u$; the subspace $W_u \oplus \{w'\}$ is mapped by B into the lower-dimensional W_u, and B is singular.

Now the ratio of the number of nonsingular B's commuting with A to the total number $p^{\alpha M}$ of matrices commuting with A is the same as the ratio of the number of nonsingular $\tilde{B}$'s to the total number. Since

$$\tilde{B} = \tilde{B}_1 \oplus \cdots \oplus \tilde{B}_n$$

is nonsingular if and only if each $\tilde{B}_u$ is so, this latter ratio is

$$f(b_1)f(b_2)\cdots f(b_n).$$

Hence

$$\mu = x^{-M}f(b_1)f(b_2)\cdots f(b_n),$$

and

$$g(\pi) = \frac{\nu}{\mu} = \frac{x^{-n^2}f(n)}{x^{-M}f(b_1)f(b_2)\cdots f(b_n)}.$$

The probability of nilpotence is therefore given by

$$P = f(n)\sum_{\pi}\frac{x^{s_1^2+s_2^2+\cdots+s_n^2}}{f(b_1)f(b_2)\cdots f(b_n)}.$$

The final stage in the proof is to establish the identity given in the following lemma:[4]

LEMMA.

$$\frac{x^n}{f(n)} = \sum_{\pi}\frac{x^{s_1^2+s_2^2+\cdots+s_n^2}}{f(b_1)\cdots f(b_n)}, \tag{4}$$

the summation being over all partitions

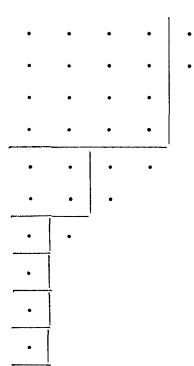
$$\pi : n = b_1 n + b_2(n-1) + \cdots + b_n\cdot 1, \tag{5}$$

where $b_u \geqq 0$, *and*

$$s_u = b_1 + b_2 + \cdots + b_u\,.$$

Proof. The left-hand side of (4) is the generating function for the number of partitions of an integer N into exactly n parts. With each such partition π^*, we associate a partition π of n as follows. Exhibit π^* as a graph, with the parts in decreasing order represented by horizontal lines of nodes, the left-hand nodes of all the parts being arranged in a vertical line. For example, the partition $30 = 5+5+4+4+4+3+2+1+1+1$ of $N = 30$ into $n = 10$ parts would have the graph

(π^*)



[4] For background material on partitions, see G. H. HARDY AND E. M. WRIGHT, *An introduction to the theory of numbers*, Oxford, 1938.

We denote by s_n the side of the largest square in the upper left corner of π^* (the Durfee square). In the example, $s_{10} = 4$, and the square is indicated by the lines. Removing the first s_n parts from π^*, we have left another partition $(4 + 3 + 2 + 1 + 1 + 1 = 12)$. Denote by s_{n-1} the side of the Durfee square for this partition ($s_9 = 2$). Remove the next s_{n-1} parts to get a third partition $(2 + 1 + 1 + 1 = 5)$ and form its Durfee square, of side s_{n-2} ($s_8 = 1$). Continuing in this way, we obtain the nonincreasing sequence $s_n \geqq s_{n-1} \geqq s_{n-2} \geqq \cdots \geqq s_1 \geqq 0$. (In our example, $s_{10} = 4$, $s_9 = 2$, $s_8 = s_7 = s_6 = s_5 = 1$, $s_4 = s_3 = s_2 = s_1 = 0$.) Clearly

$$n = s_1 + s_2 + \cdots + s_n .$$

Define $b_u = s_u - s_{u-1} \geqq 0$ $(u = 1, 2, \cdots, n)$, with $s_0 = 0$. Then if we use the relation

$$s_u = b_1 + b_2 + \cdots + b_u ,$$

we have

$$n = b_1 n + b_2(n - 1) + \cdots + b_n \cdot 1.$$

Thus with each partition π^* of an integer N into exactly n parts is associated a certain partition π of n given by the process just described. In our example, $b_{10} = 2, b_9 = 1, b_8 = b_7 = b_6 = 0, b_5 = 1, b_4 = b_3 = b_2 = b_1 = 0$, and π is given by

$$10 = 0\cdot 10 + 0\cdot 9 + 0\cdot 8 + 0\cdot 7 + 1\cdot 6 + 0\cdot 5 + 0\cdot 4 + 0\cdot 3 + 1\cdot 2 + 2\cdot 1,$$

or, in more customary form,

$$10 = 6 + 2 + 1 + 1.$$

For a given π, it is possible to reconstruct partially the original π^* by setting down in order the Durfee squares of sides $s_n, \cdots, s_1$, the total content being $M = s_1^2 + s_2^2 + \cdots + s_n^2$. To complete the reconstruction, we require the residual partitions $\pi_n, \cdots, \pi_1$ which lie to the right of the corresponding squares, with total content $N - M$. In our example, π_{10} is $2 = 1 + 1$, π_9 is $3 = 2 + 1$, π_8 is $1 = 1$, and all the others are vacuous. These residual partitions are restricted by the following conditions:

(n) π_n has at most s_n parts,
($n - 1$) π_{n-1} has at most s_{n-1} parts, of size at most $s_n - s_{n-1} = b_n$,
($n - 2$) π_{n-2} has at most s_{n-2} parts, of size at most $s_{n-1} - s_{n-2} = b_{n-1}$,
. . . .
(2) π_2 has at most s_2 parts, of size at most $s_3 - s_2 = b_3$,
(1) π_1 has at most s_1 parts, of size at most $s_2 - s_1 = b_2$,

and by the overall condition that the total content is $N - M$. If the content of π_j is C_j, then the number of partitions π_n satisfying condition (n) is the coefficient of x^{C_n} in

$$\frac{1}{(1 - x)(1 - x^2) \cdots (1 - x^{s_n})} = \frac{1}{f(s_n)} .$$

For $j < n$, the number of partitions π_j satisfying condition (j) is the coefficient[5] of x^{C_j} in

$$\frac{f(s_j + b_{j+1})}{f(s_j)f(b_{j+1})} = \frac{f(s_{j+1})}{f(s_j)f(b_{j+1})}.$$

Since the conditions (n) to (1) are independent, the total number of sets $(\pi_n, \cdots, \pi_1)$ for which $C_1 + C_2 + \cdots + C_n = N - M$ is the coefficient of x^{N-M} in

$$\frac{1}{f(s_n)} \cdot \frac{f(s_n)}{f(b_n)f(s_{n-1})} \cdot \frac{f(s_{n-1})}{f(b_{n-1})f(s_{n-2})} \cdots \frac{f(s_2)}{f(b_2)f(s_1)} = \frac{1}{f(b_n)f(b_{n-1}) \cdots f(b_2)f(b_1)}$$

since $s_1 = b_1$. This is the same as the coefficient of x^N in

$$\frac{x^M}{f(b_1) \cdots f(b_n)}.$$

This represents the contribution of the particular partition π to the total number of π^*. Summing over all π, we get the right side of (4). This completes the proof.

THEOREM 2. *The probability that an n by n matrix over the integers* mod m *be nilpotent is $(p_1 p_2 \cdots p_k)^{-n}$, where $p_1, \cdots, p_k$ are the distinct prime factors of m.*

Proof. Let $P(m)$ denote the required probability. If $(m_1, m_2) = 1$, then $P(m_1 m_2) = P(m_1)P(m_2)$, since a matrix is nilpotent mod $m_1 m_2$ if and only if it is nilpotent mod m_1 and m_2, and these events are independent. Thus it is sufficient to prove the theorem for $m = p^\beta$, where p is a prime. By Theorem 1, we may assume that $\beta > 1$.

Let A be an arbitrary matrix with elements satisfying $0 \leqq a_{ij} < p^\beta$. Then we may write, uniquely,

$$A = B + pC,$$

where $0 \leqq b_{ij} < p$, $0 \leqq c_{ij} < p^{\beta-1}$. It is easily verified that A is nilpotent mod p^β if and only if B is nilpotent mod p. Hence $P(p^\beta) = P(p) = p^{-n}$, and the theorem is proved.

It is clear that the result can easily be extended to analogous results for matrices over finite commutative rings and to similar situations.

UNIVERSITY OF PENNSYLVANIA
PHILADELPHIA, PENNSYLVANIA

[5] See, for example, P. A. MAC MAHON, *Combinatory analysis*, Cambridge, 1916, vol. 2, p. 5.