Nilpotent Matrices

Eytan Chong

1 Introduction

Let n and k be non-negative integers, and define $\mathcal{A}(n,k)$ to be the set of all $n \times n$ $(0,1)$-matrices containing exactly k ones that square to $\mathbf{0}$. In this note, we investigate the function $f(n,k) = k!\,|\mathcal{A}(n,k)|$.

		k									
	$f(n,k)$	0	1	2	3	4	5	6	7	8	9
n	0	1									
	1	1									
	2	1	2								
	3	1	6	6							
	4	1	12	36	32	6					
	5	1	20	120	280	280	120	20			
	6	1	30	300	1320	2910	3492	2400	960	210	20

Table 1: $|\mathcal{A}(n,k)|$ for small values of n and k. (A052296)

2 Preliminaries

2.1 Graph Theory

We begin with some several elementary definitions in graph theory.

Definition 1 (Adjacency Matrices)**.**

- Let (V,E) be a *directed graph* whose vertices are indexed as $V = \{v_1, \ldots, v_n\}$. Its adjacency matrix is the $n \times n$ matrix $\mathbf{M}$ with entries given by

$$\mathbf{M}_{ij} = \begin{cases} 1, & \text{if } (v_i, v_j) \in E, \\ 0, & \text{otherwise,} \end{cases}$$

  where $i, j \in [n]$.

- Let (U,W,E) be a *directed bipartite graph* whose partite sets are indexed as $U = \{u_1, \ldots, u_m\}$ and $W = \{w_1, \ldots, w_n\}$. Its adjacency matrix is the $m \times n$ matrix $\mathbf{M}$ with entries given by

$$\mathbf{M}_{ij} = \begin{cases} 1, & \text{if } (u_i, w_j) \in E, \\ 0, & \text{otherwise,} \end{cases}$$

  where $i \in [m]$ and $j \in [n]$.

Date: February 14, 2026
Latest revision: https://asdia.dev/expository/nilpotent-matrices.pdf

Definition 2. A *directed walk of length n* is a sequence of vertices $v_1, \ldots, v_{n+1}$ (possibly with repetition) such that there exists a directed edge (v_i, v_{i+1}) for all $1 \le i \le n$.

It is well-known that powers of an adjacency matrix are closely related to directed walks.

Proposition 3. *Let $\mathbf{M}$ be the adjacency matrix of a directed graph whose vertices are indexed as $V = \{v_1, \ldots, v_n\}$. Then $(\mathbf{M}^k)_{ij}$ counts the number of directed walks of length k from v_i to v_j.*

2.2 Combinatorics

The *Stirling numbers of the first and second kind* are written as $\left[{n \atop k}\right]$ and $\left\{{n \atop k}\right\}$, respectively. The *falling factorial* is written as $(x)_n = x(x-1)\ldots(x-n+1)$. We write $[n]$ to mean the set $\{1, \ldots, n\}$.

Throughout this paper, we make use of the following well-known combinatorial identities.

Proposition 4 ([2, 5])**.** *Let n and k be non-negative integers with $n \ge k$, and x a non-negative real number.*

(4.1) $$\binom{n}{k} = \frac{(n)_k}{k!},$$

(4.2) $$\left\{{n \atop k}\right\} = \frac{1}{k!}\sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i}i^n,$$

(4.3) $$(x)_n = \sum_{k=0}^{n}(-1)^{n-k}\left[{n \atop k}\right]x^n,$$

(4.4) $$\sum_{k=0}^{n}\left\{{n \atop k}\right\}(x)_k = x^n.$$

Definition 5. The *Stirling transform* of a sequence $\{a_n\}_{n\ge0}$, denoted $\mathcal{S}[\{a_n\}_{n\ge0}]$, is defined to be the sequence $\{b_n\}_{n\ge0}$ given by

$$b_n = \sum_{k=0}^{\infty}\left\{{n \atop k}\right\}a_k.$$

The inverse transform, denoted $\mathcal{S}^{-1}[\{b_n\}_{n\ge0}]$, is

$$a_n = \sum_{k=0}^{\infty}(-1)^{n-k}\left[{n \atop k}\right]b_k.$$

Example 6. From Identity (4.4), we see that $\mathcal{S}[\{(x)_n\}_{n\ge0}] = \{x^n\}_{n\ge0}$. Conversely, Identity (4.3) tells us that $\mathcal{S}^{-1}[\{x^n\}_{n\ge0}] = \{(x)_n\}_{n\ge0}$.

Proposition 7 ([1])**.** *Suppose $\mathcal{S}[\{a_n\}_{n\ge0}] = \{b_n\}_{n\ge0}$ and define the exponential generating functions*

$$A(x) = \sum_{n=0}^{\infty}\frac{a_n}{n!}x^n \quad \textit{and} \quad B(x) = \sum_{n=0}^{\infty}\frac{b_n}{n!}x^n.$$

Then $B(x) = A(\mathrm{e}^x - 1)$.

3 Properties of $f(n,k)$

3.1 Generating Function

Theorem A. *Let $F(x;n) = \sum_{k=0}^{\infty} f(n,k)\, x^k/k!$ be the exponential generating function of the sequence $\{f(n,k)\}_{k\ge0}$. Then*

$$F(x;n) = \sum_{i=0}^{n}\binom{n}{i}\left((1+x)^{n-i}-1\right)^i.$$

This result appears in [3], but a proof was omitted. We provide a proof here for completeness.

Proof. Note that

$$F(x;n) = \sum_{k=0}^{\infty} \frac{f(n,k)}{k!} x^k = \sum_{k=0}^{\infty} |\mathcal{A}(n,k)|\, x^k.$$

We thus find the ordinary generating function for $|\mathcal{A}(n,k)|$.

Let $\mathbf{A} \in \bigcup_{k\geq 0} \mathcal{A}(n,k)$. If $\mathbf{A}_{ij} = 1$, then $\mathbf{A}_{jm} = 0$ for all $m \in [n]$; if $\mathbf{A}_{jm} = 1$ for some $m \in [n]$, we obtain the contradiction

$$\left(\mathbf{A}^2\right)_{im} = \sum_{t=1}^{n} \mathbf{A}_{it}\mathbf{A}_{tm} \geq \mathbf{A}_{ij}\mathbf{A}_{jm} = 1 \neq 0.$$

That is to say, if column j has a 1, then row j must be all zeroes.

Given two complementary sets A and $B = [n] \setminus A$, we let A be the indices of columns that *must* have a 1, and B the indices of rows that *could* have a 1. We now count the number of ways to construct $\mathbf{A}$ given some choice of A. Within each column $j \in A$, we can only place 1's in the rows indexed by B. Since the column cannot be all zeroes, this amounts to choosing a non-empty subset of B. Hence, the generating function recording all possibilities for that particular column j is

$$\sum_{i=1}^{|B|} \binom{|B|}{i} x^i = (1+x)^{|B|} - 1.$$

As the choices for different columns in A are independent, the ordinary generating function over all columns is

$$\left((1+x)^{|B|} - 1\right)^{|A|}.$$

Fix $|A| = i$, so $|B| = n - i$. There are $\binom{n}{i}$ choices for A. Summing over all possible sizes i, we finally obtain the generating function

$$F(x;n) = \sum_{i=0}^{n} \binom{n}{i} \left((1+x)^{n-i} - 1\right)^i.$$

□

Corollary 8. *Let* $S(n) = \sum_{k=0}^{n} |\mathcal{A}(n,k)|$ *be the total number of* $n \times n$ *matrices that square to* $\mathbf{0}$. *Then*

$$S(n) = \sum_{i=0}^{n} \binom{n}{i} \left(2^{n-i} - 1\right)^i.$$

n	0	1	2	3	4	5	6
$S(n)$	1	1	3	13	87	841	11643

Table 2: $S(n)$ for small values of n. (A001831)

3.2 Formulas

In this subsection, we give several formulas for $f(n,k)$. We begin with the following combinatorial lemma.

Lemma 9. *The number of directed bipartite graphs (S,T,E) with exactly k edges and no isolated vertices is given by*

$$\frac{s!\,t!}{k!}\sum_{i=0}^{k}(-1)^{k-i}\begin{bmatrix}k\\i\end{bmatrix}\begin{Bmatrix}i\\s\end{Bmatrix}\begin{Bmatrix}i\\t\end{Bmatrix},$$

where $s=|S|$ and $t=|T|$.

Proof. By considering the adjacency matrix, there is a one-to-one correspondence between directed bipartite graphs with exactly k edges and no isolated vertices, and $s\times t$ $(0,1)$-matrices with exactly k 1's and no all-zero rows or columns.

Let # denote the number of such matrices (equivalently, the number of graphs). We count # using inclusion-exclusion. Let R_i and C_j denote the events that the ith row and jth column, respectively, are all zero. Then the number of matrices with no all-zero rows or columns is

$$\# = \sum_{A\subseteq S}\sum_{B\subseteq T}(-1)^{|A|+|B|}\left|\bigcap_{a\in A}R_a\cap\bigcap_{b\in B}C_b\right|.$$

Grouping the terms by the sizes of A and B gives

$$\# = \sum_{i=0}^{s}\sum_{j=0}^{t}\sum_{\substack{A\subseteq S\\|A|=i}}\sum_{\substack{B\subseteq T\\|B|=j}}(-1)^{i+j}\left|\bigcap_{a\in A}R_a\cap\bigcap_{b\in B}C_b\right|.$$

Fix $|A|=i$ and $|B|=j$. There are $\binom{s}{i}$ choices for A and $\binom{t}{j}$ choices for B. Consider the event $\bigcap_{a\in A}R_a\cap\bigcap_{b\in B}C_b$, in which the rows indexed by A and columns indexed by B are all zero. There are $(s-i)(t-j)$ remaining positions to place the k 1's in, so

$$\left|\bigcap_{a\in A}R_a\cap\bigcap_{b\in B}C_b\right| = \binom{(s-i)(t-j)}{k}.$$

Substituting these values into our expression for # yields

$$\# = \sum_{i=0}^{s}\sum_{j=0}^{t}(-1)^{i+j}\binom{s}{i}\binom{t}{j}\binom{(s-i)(t-j)}{k}.$$

Applying the combinatorial identities listed in Proposition 4 finishes the proof.

$$
\begin{aligned}
\# &= \sum_{i=0}^{s}\sum_{j=0}^{t}(-1)^{i+j}\binom{s}{i}\binom{t}{j}\binom{(s-i)(t-j)}{k} \\
&\overset{(4.1)}{=} \sum_{i=0}^{s}\sum_{j=0}^{t}(-1)^{i+j}\binom{s}{i}\binom{t}{j}\frac{((s-i)(t-j))_k}{k!} \\
&\overset{(4.3)}{=} \frac{1}{k!}\sum_{i=0}^{s}\sum_{j=0}^{t}(-1)^{i+j}\binom{s}{i}\binom{t}{j}\sum_{m=0}^{k}(-1)^{k-m}\begin{bmatrix}k\\m\end{bmatrix}(s-i)^m(t-j)^m \\
&= \frac{s!\,t!}{k!}\sum_{m=0}^{k}(-1)^{k-m}\begin{bmatrix}k\\m\end{bmatrix}\left[\frac{1}{s!}\sum_{i=0}^{s}(-1)^i\binom{s}{i}(s-i)^m\right]\left[\frac{1}{t!}\sum_{j=0}^{t}(-1)^j\binom{t}{j}(t-j)^m\right] \\
&\overset{(4.4)}{=} \frac{s!\,t!}{k!}\sum_{m=0}^{k}(-1)^{k-m}\begin{bmatrix}k\\m\end{bmatrix}\begin{Bmatrix}m\\s\end{Bmatrix}\begin{Bmatrix}m\\t\end{Bmatrix}.
\end{aligned}
$$

□

With this result, we are ready to find a formula for $f(n,k)$.

Theorem B.

$$
f(n,k) = \sum_{i=0}^{k}(-1)^{k-i}\begin{bmatrix}k\\i\end{bmatrix}\sum_{0\le s,t\le i}\begin{Bmatrix}i\\s\end{Bmatrix}\begin{Bmatrix}i\\t\end{Bmatrix}(n)_{s+t}.
$$

Proof. By Proposition 3, there is a one-to-one correspondence between $\mathcal{A}(n,k)$ and the set of all directed graphs (V,E) with $|V| = n$ and $|E| = k$ that do not contain any directed walks of length 2. We count the number of such graphs.

Let S and T be the sets of vertices with non-zero outdegree and indegree respectively. Since all walks are of length one, the sets S and T are disjoint. Thus, for fixed sizes $s = |S|$ and $t = |T|$, there are $\binom{n}{s}\binom{n-s}{t}$ ways to choose S and T from V. Next, Lemma 9 tells us that for any choice of S and T, there are

$$
\frac{s!\,t!}{k!}\sum_{i=0}^{k}(-1)^{k-i}\begin{bmatrix}k\\i\end{bmatrix}\begin{Bmatrix}i\\s\end{Bmatrix}\begin{Bmatrix}i\\t\end{Bmatrix}
$$

ways to draw k edges from vertices in S to vertices in T. Enumerating over all possible sizes s and t, the number of directed graphs (and thus matrices) is

$$
|\mathcal{A}(n,k)| = \sum_{s=0}^{k}\sum_{t=0}^{k}\binom{n}{s}\binom{n-s}{t}\frac{s!\,t!}{k!}\sum_{i=0}^{k}(-1)^{k-i}\begin{bmatrix}k\\i\end{bmatrix}\begin{Bmatrix}i\\s\end{Bmatrix}\begin{Bmatrix}i\\t\end{Bmatrix}.
$$

Writing $\binom{n}{s}\binom{n-s}{t}$ as $(n)_{s+t}/s!\,t!$ and noting that the summand vanishes when $s,t > i$, we obtain

$$
f(n,k) = k!\,|\mathcal{A}(n,k)| = \sum_{i=0}^{k}(-1)^{k-i}\begin{bmatrix}k\\i\end{bmatrix}\sum_{0\le s,t\le i}\begin{Bmatrix}i\\s\end{Bmatrix}\begin{Bmatrix}i\\t\end{Bmatrix}(n)_{s+t}.
$$

□

Remark. Theorem B generalizes the $k = 2$ case as discussed by the author in [7].

Applying Identity (4.4) to the formula given by Theorem B, the triple sum collapses to the following double sum expression for $f(n,k)$.

Proposition 10.

$$f(n,k) = \sum_{i=0}^{k} (-1)^{k-i} \begin{bmatrix} k \\ i \end{bmatrix} \sum_{s=0}^{i} \begin{Bmatrix} i \\ s \end{Bmatrix} (n)_s (n-s)^i.$$

Proof. Write $(n)_{s+t} = (n)_s(n-s)_t$. By Theorem B, this gives

$$f(n,k) = \sum_{i=0}^{k} (-1)^{k-i} \begin{bmatrix} k \\ i \end{bmatrix} \sum_{s=0}^{i} \begin{Bmatrix} i \\ s \end{Bmatrix} (n)_s \sum_{t=0}^{i} \begin{Bmatrix} i \\ t \end{Bmatrix} (n-s)_t$$

$$\overset{(4.4)}{=} \sum_{i=0}^{k} (-1)^{k-i} \begin{bmatrix} k \\ i \end{bmatrix} \sum_{s=0}^{i} \begin{Bmatrix} i \\ s \end{Bmatrix} (n)_s (n-s)^i.$$

□

Next, we use Theorem B to write $f(n,k)$ as a polynomial in n, with coefficients depending on k.

Proposition 11.

$$f(n,k) = \sum_{p=0}^{2k} \left[(-1)^p \sum_{i=\lceil p/2 \rceil}^{k} (-1)^{k-i} \begin{bmatrix} k \\ i \end{bmatrix} \sum_{\substack{0 \le s,t \le i \\ s+t \ge p}} (-1)^{s+t} \begin{Bmatrix} i \\ s \end{Bmatrix} \begin{Bmatrix} i \\ t \end{Bmatrix} \begin{bmatrix} s+t \\ p \end{bmatrix} \right] n^p.$$

Proof. From Theorem B, we have that

$$f(n,k) = \sum_{i=0}^{k} (-1)^{k-i} \begin{bmatrix} k \\ i \end{bmatrix} \sum_{s=0}^{i} \begin{Bmatrix} i \\ s \end{Bmatrix} (n)_s (n-s)^i.$$

Using Identity (4.3) to expand $(n)_{s+t}$ as a sum of monomials, we get

$$f(n,k) = \sum_{i=0}^{k} (-1)^{k-i} \begin{bmatrix} k \\ i \end{bmatrix} \sum_{0 \le s,t \le i} \begin{Bmatrix} i \\ s \end{Bmatrix} \begin{Bmatrix} i \\ t \end{Bmatrix} \sum_{p=0}^{s+t} (-1)^{s+t-p} \begin{bmatrix} s+t \\ p \end{bmatrix} n^p.$$

Interchanging the order of summation so that we sum over p first yields the desired expression. □

3.3 Stirling Transform and Chromatic Polynomials

Let $P(K_{k,k}, n)$ be the chromatic polynomial of the complete bipartite graph $K_{k,k}$. [6] gives the closed-form

$$P(K_{k,k}, n) = \sum_{0 \le s,t \le k} \begin{Bmatrix} k \\ s \end{Bmatrix} \begin{Bmatrix} k \\ t \end{Bmatrix} (n)_{s+t}.$$

It follows from Theorem B that $\{P(K_{k,k}, n)\}_{k \ge 0}$ is the Stirling transform of $\{f(n,k)\}_{k \ge 0}$.

Proposition 12. *$P(K_{k,k}, n)$ has exponential generating function*

$$\sum_{i=0}^{n} \binom{n}{i} \left(e^{xi} - 1\right)^{n-i}.$$

Proof. Let $G(x;n)$ be the exponential generating function of $\{P(K_{k,k}, n)\}_{k \ge 0}$. Since $\{P(K_{k,k}, n)\}_{k \ge 0}$ is the Stirling transform of $\{f(n,k)\}_{k \ge 0}$, we have by Proposition 7 that $G(x;n) = F(e^x - 1; n)$. Using the formula for $F(x;n)$ given in Theorem A, we see that

$$G(x;n) = \sum_{i=0}^{n} \binom{n}{i} \left(e^{xi} - 1\right)^{n-i}.$$

□

3.4 Factorization

Proposition 13. *For fixed n, the support of $f(n,k)$ is $k = 0, \dots, \lfloor n^2/4 \rfloor$.*

Proof. The degree of the ith term of $F(x;n)$ is $i(n-i)$, which attains a maximum of $\lfloor n^2/4 \rfloor$ when $i = \lfloor n/2 \rfloor$. Thus, $f(n,k) = 0$ for all $k > \lfloor n^2/4 \rfloor$. To show that $f(n,k)$ is non-zero for $k = 0, \dots, \lfloor n^2/4 \rfloor$, it suffices to construct a matrix $\mathbf{A} \in \mathcal{A}(n, \lfloor n^2/4 \rfloor)$, since for any $k < \lfloor n^2/4 \rfloor$ a matrix in $\mathcal{A}(n,k)$ may be obtained by replacing some of the entries of $\mathbf{A}$ from 1 to 0.

Let $\mathbf{0}_{a \times b}$ and $\mathbf{1}_{a \times b}$ denote the $a \times b$ blocks whose entries all 0 and 1, respectively. We construct $\mathbf{A}$ according to the parity of n.

Case 1: *n is even*. Write $n = 2m$, so $\lfloor n^2/4 \rfloor = m^2$. Then

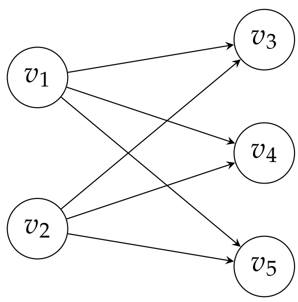
$$\mathbf{A} = \begin{pmatrix} \mathbf{0}_{m \times m} & \mathbf{1}_{m \times m} \\ \mathbf{0}_{m \times m} & \mathbf{0}_{m \times m} \end{pmatrix} \in \mathcal{A}(2m, m^2).$$

Case 2: *n is odd*. Write $n = 2m+1$, so $\lfloor n^2/4 \rfloor = m^2 + m$. Then

$$\mathbf{A} = \begin{pmatrix} \mathbf{0}_{m \times m} & \mathbf{1}_{m \times (m+1)} \\ \mathbf{0}_{(m+1) \times m} & \mathbf{0}_{(m+1) \times (m+1)} \end{pmatrix} \in \mathcal{A}(2m+1, m^2+m).$$

□

Remark. In the case where $n = 2m$ is even, $\mathbf{A}$ is the adjacency matrix of the directed complete bipartite graph $\overrightarrow{K}_{m,m}$. When $n = 2m+1$ is odd, $\mathbf{A}$ is the adjacency matrix of the directed complete bipartite graph $\overrightarrow{K}_{m,m+1}$.



$$\begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathcal{A}(5,6).$$

Figure 1: The directed complete bipartite graph $\overrightarrow{K}_{2,3}$ and its adjacency matrix.

Proposition 14. *Fix k and let $\alpha_k = \lceil 2\sqrt{k} \rceil$. Then $f(n,k) = (n)_{\alpha_k} P(n,k)$ for some monic polynomial $P(n,k) \in \mathbb{Z}[n]$ of degree $2k - \alpha_k$ with no integer roots.*

Proof. Recall from Theorem B that

$$f(n,k) = \sum_{i=0}^{k} (-1)^{k-i} \begin{bmatrix} k \\ i \end{bmatrix} \sum_{0 \le s,t \le i} \begin{Bmatrix} i \\ s \end{Bmatrix} \begin{Bmatrix} i \\ t \end{Bmatrix} (n)_{s+t}.$$

It is easy to see that $f(n,k) \in \mathbb{Z}[n]$. Next, we note that $\deg((n)_{s+t}) = s + t \le 2k$, with equality if and only if $s = t = i = k$. The corresponding term is

$$(-1)^{k-k} \begin{bmatrix} k \\ k \end{bmatrix} \begin{Bmatrix} k \\ k \end{Bmatrix} \begin{Bmatrix} k \\ k \end{Bmatrix} (n)_{k+k} = (n)_{2k},$$

which is monic. It follows that $f(n,k)$ is a monic polynomial of degree $2k$.

By Proposition 13, we see that $f(n,k)$ vanishes if and only if $n \le \lceil 2\sqrt{k} \rceil = \alpha_k$. Thus, the only integer roots of $f(n,k)$ are $n = 0, 1, \ldots, \alpha_k$, giving the factorization $f(n,k) = (n)_{\alpha_k} P(n,k)$, where $P(n,k) \in \mathbb{Z}[n]$ is a monic polynomial of degree $2k - \alpha_k$ with no integer roots. □

k	$P(n,k)$
0	1
1	1
2	$-1 + n$
3	$4 - 3n + n^2$
4	$86 - 96n + 43n^2 - 10n^3 + n^4$
5	$-810 + 886n - 415n^2 + 105n^3 - 15n^4 + n^5$
6	$-46440 + 59752n - 34168n^2 + 11341n^3 - 2380n^4 + 320n^5 - 26n^6 + n^7$

Table 3: The polynomials $P(n,k)$ for $k = 1, \ldots, 6$

References

[1] M. Bernstein and N. J. A. Sloane. Some canonical sequences of integers. *Linear Algebra and its Applications*, 226-228:57–72, 1995.

[2] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics*. Addison-Wesley, Massachusetts, second edition, 1994.

[3] G. Kilibarda and V. Jovovic. Antichains of Multisets. *Journal of Integer Sequences*, 7, 2004.

[4] OEIS Foundation Inc., 2026. The On-Line Encyclopedia of Integer Sequences. Published electronically at `https://oeis.org`.

[5] J. Quaintance and H. W. Gould. *Combinatorial Identities for Stirling Numbers*. World Scientific, Singapore, 2016.

[6] J. R. Swenson. The Chromatic Polynomial of a Complete Bipartite Graph. *American Mathematical Monthly*, 80:797–798, 1973.

[7] Various authors. Determine the probability that $A^2 = O$? Mathematics Stack Exchange. `https://math.stackexchange.com/q/4367731`. Accessed 2025-09-30.