

A MATRIX-BASED APPROACH TO COMBINATORIAL IDENTITIES INVOLVING BINOMIAL COEFFICIENTS AND STIRLING NUMBERS

ABSTRACT. In this paper, we present a novel approach to matrix representation. By employing the derivative method of a matrix defined through shift matrices, we show how the sums of specific combinatorial sequences, including binomial coefficients and Stirling numbers, can be directly computed.

Keywords: matrix, sequences, stirling, coefficients

1. INTRODUCTION

Michael Spivey effectively organizes and presents a variety of methods for working with binomial coefficients in ten chapters. The book explores techniques from algebra, calculus, linear algebra, and combinatorics, [14]. Many authors have explored various combinatorial identities involving binomial coefficients and their generalizations, presenting proofs of the resulting identities using different methods, [2, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16].

Working on known problems through matrices is often more convenient because matrices provide a systematic and structured way of organizing and solving problems. When dealing with large-scale problems, matrices provide a scalable way to handle increasing complexity. Whether it's large systems of equations, high-dimensional data, or complex transformations, matrices enable efficient management of large amounts of information. Matrices provide a general framework that can be applied to a wide range of problems. Whether you are dealing with linear algebra, optimization, machine learning, or graph theory, the principles of matrix operations remain consistent, allowing for easier generalization of techniques and results, [1, 2, 3, 10, 11, 15, 16, 17].

In [8], John G. Kemeny introduces the concept of the derivative of a matrix and presents a method for proving combinatorial identities by employing denumerably infinite matrices $M = [M_{ij}]$, where $i, j = 0, 1, 2, 3, \dots$. He then uses these matrices to express several combinatorial identities. In this paper, based on the concept introduced in [8], we showed how various sums involving binomial coefficients and Stirling numbers can be computed directly. In the next section, we present some results and examples. Afterward, we explore certain vectors and provide proofs of some combinatorial identities using this method.

2000 *Mathematics Subject Classification.* Primary 15A24, 05A19; Secondary 05A10, 11B65.

Key words and phrases. Matrix equations and identities, combinatorial identities, binomial coefficients, Stirling numbers.

2. DERIVATIVE OF MATRICES AND VECTORS

When working with infinite matrices, we must be cautious, as some laws may not hold. To avoid this issue, we ensure that the matrices we use have only a finite number of nonzero terms in the computation of each component of a product. For infinite matrices, we introduce some standard definitions. For instance, a diagonal matrix is one in which all entries outside the main diagonal are zero. A matrix is called lower triangular if all entries above the main diagonal are zero, while an upper triangular matrix has all entries below the main diagonal set to zero. The identity matrix is the matrix with ones on the main diagonal and zeros elsewhere. We will denote zero matrices and vectors by $\mathbf{0}$, and the specific type (matrix or vector) will be clear from the context.

Following the approach outlined by Kemeny [8], we define two shift matrices as follows:

$$L_{ij} = \delta_{i,j+1} \quad \text{and} \quad U = \delta_{i+1,j},$$

where $\delta_{i,j}$ is the Kronecker delta symbol.

The effects of the matrices L and U on a matrix M are described by

$$\begin{aligned} (LM)_{ij} &= M_{i-1,j} \text{ adds a null row } \mathbf{0} \\ (ML)_{ij} &= M_{i,j+1} \text{ removes column } \mathbf{0} \\ (UM)_{ij} &= M_{i+1,j} \text{ removes row } \mathbf{0} \\ (MU)_{ij} &= M_{i,j-1} \text{ adds a null column } \mathbf{0}. \end{aligned}$$

Clearly, there are many possible permutations. For instance, UML represents the matrix \widehat{M} being shifted up and left along the main diagonal. This operation is denoted by $\widehat{M} = UML$.

We also define the row vector $e = [e_j]$, where $e_j = 1$ when $j = 0$ and $e_j = 0$ otherwise, and the column vector $f = [f_i]$, where $f_i = 1$ when $i = 0$ and $f_i = 0$ otherwise. Additionally, we denote the row vector of all ones as $\mathbf{1}$.

The concept of the derivative of a matrix is introduced through the following definition.

Definition 1. *A derivative of a matrix X is defined as*

$$X' = XL - LX,$$

and the derivative of a row vector y is given by

$$y' = yL - y.$$

Using the above definitions and relations, we can express the derivatives as:

$$\begin{aligned} X'_{ij} &= X_{i,j+1} - X_{i-1,j}, \\ y' &= y_{j+1} - y_j. \end{aligned}$$

Theorem 1. *The derivative possesses the following properties*

- (i) $(cX)' = cX'$, where c is a constant,
- (ii) $(X + Y)' = X' + Y'$,
- (iii) $(XY)' = XY' + X'Y$.

Proof. The first two results can be easily proven using the definition. Now, let's prove result (iii).

$$\begin{aligned}
 (XY)' &= XYL - LXY \\
 &= XYL - XLY + XLY - LXY \\
 &= X(YL - LY) + (XL - LX)Y \\
 &= XY' + X'Y.
 \end{aligned}$$

□

Similarly, the following properties of derivatives hold for vectors.

- (i) $(cy)' = cy'$, where c is a constant,
- (ii) $(y + L)' = y' + L'$,
- (iii) $(yX)' = y'X + yX'$.

The following theorem will play a crucial role in proving the identities.

Theorem 2. *For given g, K, A_n , and B_n , if the matrices B_n are upper triangular then the conditions*

- (i) $Xf = g$
- (ii) $X' = \sum_n A_n X B_n + K$

determine a unique matrix X .

Proof. The proof proceeds by induction on the columns of X . The 0th column of X is determined by (i). Assume that the columns are uniquely determined up to column j . By the definition of the derivative, the relation $XL = X' + LX$ determines column $j + 1$. If B_n is upper triangular, the sum in (ii) involves the columns of X up to column j . Moreover, LX also involves row j . Therefore, by the induction hypothesis, column $j + 1$ is uniquely determined. This completes the proof. □

We refer to (i) and (ii) in Theorem 2 as the defining relations. The defining relations for vectors are stated in the following theorem.

Theorem 3. *For given a, A_n, b , if A_n are upper triangular then the conditions*

- (i) $yf = a$
- (ii) $y' = \sum_n y A_n + b$

determine a unique row vector y .

Example 1. *The defining relations for certain matrices and vectors are as follows*

$$\begin{array}{ll}
 If = f & I' = \mathbf{0} \\
 Uf = \mathbf{0} & U' = fe \\
 Lf = [0, 1, 0, \dots]^T & L' = \mathbf{0} \\
 ef = 1 & e' = -e \\
 \mathbf{1}f = 1 & \mathbf{1}' = \mathbf{0},
 \end{array}$$

where I is the identity matrix.

Example 2. *Let's examine the defining relations $Df = \mathbf{0}$ and $D' = L$. By using the definition of the derivative to compute a part of D from these relations, we find that D is a diagonal matrix with entries $D_{ii} = i$.*

Example 3. Let

$$Ff = f \quad F' = LDF.$$

These relations result in a diagonal matrix F with $i!$ on the main diagonal.

Example 4. Let t be a row vector with components $t_j = 2^j$. It is straightforward to observe that

$$tf = 1 \quad t' = t,$$

which implies that t is the unique row vector with 1 on the 0^{th} column and equal to its derivative.

In general, the inverse of an infinite matrix is not well-defined. However, the inverse of an upper triangular matrix with nonzero diagonal entries can be considered. While there is an algorithm to compute the inverse of finite upper triangular matrices, this method can also be applied to infinite upper triangular matrices, though it involves an infinite number of steps. For a diagonal matrix, its inverse is also a diagonal matrix, with the diagonal entries being the reciprocals of the original entries. D does not have an inverse since $Df = \mathbf{0}$. However, $\widehat{D} = D + I$ has an inverse, and the inverse of \widehat{D}_{ii} is given by $\widehat{D}_{ii}^{-1} = \frac{1}{i+1}$.

For alternating series, we require the definition of duality.

Definition 2. [8] The dual of a matrix M is

$$\overline{M}_{ij} = (-1)^{i-j} M_{ij},$$

and the dual of a row vector r and a column vector h are

$$\overline{r}_j = (-1)^j r_j \quad \text{and} \quad \overline{h}_i = (-1)^i h_i.$$

Example 5. Let us see the properties of the dual binomial coefficients \overline{B} . Applying above definition, we get

$$\begin{aligned} \overline{B}f &= f \\ \overline{B}' &= -\overline{B}. \end{aligned}$$

Since $(\mathbf{1}\overline{B})f = 1$, $(\mathbf{1}\overline{B})' = -\mathbf{1}\overline{B}$ and $ef = 1$, $e' = -e$, we obtain that $\mathbf{1}\overline{B} = e$.

3. APPLICATIONS: MATRIX REPRESENTATION OF CERTAIN IDENTITIES

In this section, we prove several combinatorial sum identities presented in [13] using the method introduced earlier.

Identity 1. $\sum_{k=0}^n \binom{n}{k} k = n2^{n-1}$.

Proof. The sum of the n^{th} column of DB is $\mathbf{1}DB$, and the defining relations of $\mathbf{1}DB$ are as follows

$$(\mathbf{1}DB)f = 0 \quad \text{and} \quad (\mathbf{1}DB)' = \mathbf{1}DB + t.$$

Now, we will examine the row vector tUD . Upon computing a segment of it, we get the following

$$(tUD)f = 0 \quad \text{and} \quad (tUD)' = tUD + t.$$

By uniqueness we conclude that, $\mathbf{1}DB = tUD$ and this states that

$$\sum_{k=0}^n \binom{n}{k} k = n2^{n-1}.$$

□

We can generalize this result by considering the matrix $D^m B$ instead of DB , where

$$D^m := D(D - I)(D - 2I) \dots (D - (m - 1)I).$$

Identity 2. If $m \geq 1$,

$$\sum_{k=0}^n \binom{n}{k} k^{\underline{m}} = n^{\underline{m}} 2^{n-m},$$

where $k^{\underline{m}}$ is the falling factorial, $k^{\underline{m}} := k(k-1)(k-2) \dots (k-m+1)$.

Proof. The sum of the n^{th} column of $D^m B$ is $\mathbf{1} D^m B$. We will now show that the defining relations of $\mathbf{1} D^m B$ are

$$(tU^m D^m)f = 0 \quad \text{and} \quad (tU^m D^m)' = tU^m D^m + mtU^{m-1} D^{m-1}$$

and the defining relations of $tU^m D^m$ are

$$(tU^m D^m)f = 0 \quad \text{and} \quad (tU^m D^m)' = tU^m D^m + mtU^{m-1} D^{m-1}.$$

The proof will proceed by induction on m . For $m = 1$, we obtain Identity 1. Assume the result is true for $m = k$. Since the sum of the 0^{th} column of $D^k B$ is 0, multiplying the 0^{th} column by $-k$ will also yield 0. Thus,

$$(\mathbf{1} D^{k+1} B)f = (\mathbf{1} D(D - I) \dots (D - kI) B)f = 0.$$

Similarly, we can see that

$$(tU^{k+1} D^{k+1})f = 0.$$

Now, let's compute the derivatives:

$$(\mathbf{1} D^{k+1} B)' = \mathbf{1} D^{k+1} B + (k+1) \mathbf{1} D^k B$$

and

$$(tU^{k+1} D^{k+1})' = tU^{k+1} D^{k+1} + (k+1)tU^k D^k.$$

By induction hypothesis, $\mathbf{1} D^k B = tU^k D^k$. Therefore $\mathbf{1} D^{k+1} B$ and $tU^{k+1} D^{k+1}$ satisfy the same defining relations. From Theorem 3, it follows that $\mathbf{1} D^m B = tU^m D^m$. □

Identity 3.

$$\sum_{k=0}^n (-1)^k \binom{n}{k} k = -[n = 1],$$

where $[]$ is the Iverson's notation, for more details see [5].

Proof. To prove this identity, we can utilize the dual binomial coefficients. The sum of the n^{th} column of $D\bar{B}$ is $-\mathbf{1} D\bar{B}$, and the defining relations of $-\mathbf{1} D\bar{B}$ are

$$(-\mathbf{1} D\bar{B})f = 0 \quad \text{and} \quad (-\mathbf{1} D\bar{B})' = \mathbf{1} D\bar{B} - e.$$

Now, consider the vector $e\bar{U}$. The corresponding defining relations are

$$(e\bar{U}F)f = 0 \quad \text{and} \quad (e\bar{U}F)' = -e\bar{U}F - e.$$

Thus, by uniqueness, we obtain $-\mathbf{1} D\bar{B} = e\bar{U}F$, which gives the identity. □

The following identity generalizes Identity 3. We will use the matrix D^m instead of D .

Identity 4. If $m \geq 1$,

$$\sum_{k=0}^n (-1)^k \binom{n}{k} k^{\underline{m}} = (-1)^m m! [n = m].$$

Proof. For the left-hand side, we can take the vector $(-1)^m \mathbf{1} D^{\underline{m}} \overline{B}$, and for the right-hand side, we take $e \overline{U}^m F$. We can prove this identity by uniqueness and induction on m . For $m = 1$, we recover Identity 3. Assume that the identity holds for $m = k$, i.e., $(-1)^k \mathbf{1} D^{\underline{k}} \overline{B} = e \overline{U}^k F$.

For $m = k + 1$, the defining relations are:

$$\begin{aligned} ((-1)^{k+1} \mathbf{1} D^{\underline{k+1}} \overline{B}) f &= 0, \\ ((-1)^{k+1} \mathbf{1} D^{\underline{k+1}} \overline{B})' &= (-1)^k \mathbf{1} D^{\underline{k+1}} \overline{B} + (-1)^{k+1} (k+1) \mathbf{1} D^{\underline{k}} \overline{B} \end{aligned}$$

and

$$\begin{aligned} (e \overline{U}^{k+1} F) f &= 0, \\ (e \overline{U}^{k+1} F)' &= -e \overline{U}^{k+1} F - (k+1) e \overline{U}^k F. \end{aligned}$$

By the induction hypothesis, $(-1)^k \mathbf{1} D^{\underline{k}} \overline{B} = e \overline{U}^k F$, which shows that the vectors $(-1)^{k+1} \mathbf{1} D^{\underline{k+1}} \overline{B}$ and $e \overline{U}^{k+1} F$ satisfy the same defining relations. Therefore, the identity is established by induction. \square

Identity 5.

$$\sum_k \binom{n}{2k} = 2^{n-1} + \frac{1}{2} [n = 0].$$

Proof. To prove this identity, let's define the matrix $A = [a_{ij}]$, where $a_{ij} = 1$ if j is an even integer and $i = \frac{j}{2}$, and $a_{ij} = 0$ otherwise. Then, we have

$$\begin{aligned} (\mathbf{1} A B) f &= 1 \quad \text{and} \quad (\mathbf{1} A B)' = \mathbf{1} A B - e, \\ (tU + e) f &= 1 \quad \text{and} \quad (tU + e)' = tU. \end{aligned}$$

Thus, by uniqueness,

$$\mathbf{1} A B = tU + e.$$

\square

Identity 6.

$$\sum_k \binom{n}{2k+1} = 2^{n-1} - \frac{1}{2} [n = 0].$$

Proof. By using matrix AU instead of A , we obtain the result as follows

$$\mathbf{1} A U B = tU.$$

\square

The following identities generalize Identity 5 and Identity 6, respectively.

Identity 7. If $m \geq 1$, then

$$(3.1) \quad \sum_k \binom{n}{2k} k^{\underline{m}} = n(n-m-1)^{\underline{m-1}} 2^{n-2m-1} [n \geq m+1].$$

Proof. To prove this identity, we will show that

$$\mathbf{1}D^m AB = tU^{2m+1}D(D - (m+1)I)^{\underline{m-1}}.$$

The proof can be done by induction on m . For $m = 1$, we have the following relations:

$$(\mathbf{1}DAB)f = 0 \quad \text{and} \quad (\mathbf{1}DAB)' = \mathbf{1}DAB + tU(I - U),$$

$$(tU^3D)f = 0 \quad \text{and} \quad (tU^3D)' = tU^3D + tU(I - U).$$

Thus, the result is valid for $m = 1$. Now, assume that it is true for $m = k$. For $m = k + 1$, the defining relations are:

$$\begin{aligned} (\mathbf{1}D^{k+1}AB)f &= 0, \\ (\mathbf{1}D^{k+1}AB)' &= \mathbf{1}D^{k+1}AB + (k+1)\mathbf{1}D^kABU, \end{aligned}$$

and

$$\begin{aligned} (tU^{2(k+1)+1}D(D - (k+2)I)^{\underline{k}})f &= 0, \\ (tU^{2(k+1)+1}D(D - (k+2)I)^{\underline{k}})' &= tU^{2(k+1)+1}D(D - (k+2)I)^{\underline{k}} + \\ &\quad (k+1)tU^{2k+1}D(D - (k+1)I)^{\underline{k-1}}U. \end{aligned}$$

By the induction hypothesis, $\mathbf{1}D^kAB = tU^{2k+1}D(D - (k+1)I)^{\underline{k-1}}$, which gives the desired result. \square

Identity 8. If $m \geq 1$, $\sum_k \binom{n}{2k+1} k^{\underline{m}} = (n - m - 1) \underline{m} 2^{n-2m-1} [n \geq m + 1]$.

Proof. Instead of the matrix A , we take the matrix AU . Similar to the proof of Identity 7, we have

$$\mathbf{1}D^m AUB = tU^{2k+1}(D - (k+1)I)^{\underline{k}}.$$

\square

Unsigned Stirling numbers of the first kind, denoted as $S(n, k)$, count the number of permutations of n elements with exactly k permutation cycles. They can also be defined recursively by the following relation:

$$S(n, k) = S(n - 1, k - 1) + (n - 1)S(n - 1, k),$$

with the initial conditions:

$$S(0, 0) = 1, \quad S(n, 0) = 0 \text{ for } n > 0, \quad S(n, n) = 1.$$

These numbers have various combinatorial interpretations and appear in many areas of mathematics, including in the study of permutations, partitions, and generating functions. We denote $S(n, k)$ by $\begin{bmatrix} n \\ k \end{bmatrix}$. Now, we will consider the identities for the unsigned Stirling numbers of the first kind.

Identity 9.

$$\sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} = n! = \begin{bmatrix} n+1 \\ 1 \end{bmatrix}.$$

Proof. The left-hand side of the identity is the sum of the n^{th} column of the unsigned Stirling numbers of the first kind. Therefore, it is equal to the vector $\mathbf{1} |K|$, where $|K|$ represents the absolute values of the entries of K . We then have the following relations:

$$\begin{aligned} (\mathbf{1} |K|)f &= 1 \quad \text{and} \quad (\mathbf{1} |K|)' = \mathbf{1} |K| D, \\ (\mathbf{1} F)f &= 1 \quad \text{and} \quad (\mathbf{1} F)' = \mathbf{1} F D. \end{aligned}$$

Thus, we conclude that $\mathbf{1} |K| = \mathbf{1} F$. \square

Identity 10.

$$\sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} k = \begin{bmatrix} n+1 \\ 2 \end{bmatrix}.$$

Proof. If we compute the vector $\mathbf{1} D |K|$, we obtain the left-hand side of the identity. To obtain the right-hand side, we need to remove the 0^{th} column of the matrix $|K|$ and then multiply by the vector eU^2 . Therefore, we have $\mathbf{1} D |K| = eU^2 |K| L = eU \widehat{|K|}$.

Let us prove this by showing that they satisfy the same defining relations. Now, we have the following relations:

$$\begin{aligned} (\mathbf{1} D |K|)f &= 0 \quad \text{and} \quad (\mathbf{1} D |K|)' = \mathbf{1} D |K| D + \mathbf{1} |K|, \\ e(eU^2 |K| L) &= 0 \quad \text{and} \quad (eU^2 |K| L)' = eU^2 |K| L D + \mathbf{1} F. \end{aligned}$$

From Identity 9, we have already shown that $\mathbf{1} |K| = \mathbf{1} F$. Hence, the result follows. \square

We can now generalize Identity 10.

Identity 11.

$$\sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix} k^m = \begin{bmatrix} n+1 \\ m+1 \end{bmatrix} m!.$$

Proof. We will demonstrate that $\mathbf{1} D^m |K| = eU^m F \widehat{|K|}$. For $m = 1$, we obtain Identity 10. Assume that the identity holds for $m = k$. For $m = k+1$, we have the following relations:

$$\begin{aligned} (\mathbf{1} D^{k+1} |K|)f &= 0 \quad \text{and} \quad (\mathbf{1} D^{k+1} |K|)' = \mathbf{1} D^{k+1} |K| D + \mathbf{1} D^k |K|, \\ (eU^{k+1} F \widehat{|K|})f &= 0 \quad \text{and} \quad (eU^{k+1} F \widehat{|K|})' = eU^{k+1} F \widehat{|K|} D + eU^k F \widehat{|K|}. \end{aligned}$$

By the induction hypothesis, $\mathbf{1} D^k |K| = eU^k F \widehat{|K|}$, which implies that $\mathbf{1} D^{k+1} |K|$ and $eU^{k+1} F \widehat{|K|}$ satisfy the same relations. Hence, by uniqueness, we conclude that

$$\mathbf{1} D^m |K| = eU^m F \widehat{|K|}.$$

\square

4. CONCLUSION

Matrices are a powerful tool in solving various mathematical problems, especially when dealing with systems of linear equations, transformations, and data manipulation. They provide a structured way to handle large sets of numbers and can simplify complex calculations. In the literature, there are numerous studies on the proofs of identities involving binomial coefficients and Stirling numbers. In this paper, proofs of various identities are provided based on the definition of the derivative of a matrix. The identities discussed are related to binomial coefficients and Stirling numbers. However, this method also facilitates the proof of many other identities beyond those considered here.

CONFLICT OF INTERESTS

We have no conflict of interest to declare.

ACKNOWLEDGEMENTS

The author would like to thank the referees for their valuable comments which helped to improve the paper.

REFERENCES

- [1] M. Barnabei, A. Brini and G. Nicoletti, Recursive matrices and umbral calculus, *Journal of Algebra*, **75** (2) (1982), 546–573.
- [2] M. Bona, *Combinatorics of Permutations* (3rd ed.), Chapman and Hall/CRC, 2022.
- [3] L. Comtet, *Advanced Combinatorics*, Reidel, Dordrecht, 1974.
- [4] L. González, A new approach for proving or generating combinatorial identities. *International Journal of Mathematical Education in Science and Technology*, **41** (3) (2010), 359–372.
- [5] R. L. Graham, D. E. Knuth and O. Patashnik, *Concrete Mathematics*, 2nd Edition, Addison-Wesley, Reading, MA, 1994.
- [6] R. Jay Hendel, A method for uniformly proving a family of identities, *The Fibonacci Quarterly*, **60** (2) (2022), 151–163.
- [7] A. Horváth, Computer algebra proof of some identities involving binomial coefficients, *Procedia Manufacturing*, **22** (2018), 1043–1050.
- [8] J. G. Kemeny, Matrix representation for combinatorics, *Journal of Combinatorial Theory Series A*, **36** (1984), 279–306.
- [9] P. Paule, M. Schorn, A Mathematica version of Zeilberger’s algorithm for proving binomial coefficient identities, *Journal of symbolic computation*, **20**(5-6) (1995), 673–698.
- [10] F. S. Roberts, B. Tesman, B., *Applied combinatorics*, CRC Press, 2024.
- [11] G-C. Rota, *Finite Operator Calculus*, Academic Press, New York, 1975.
- [12] M. Shattuck, T. Waldhauser, Proofs of some binomial identities using the method of last squares, *The Fibonacci Quarterly*, **48** (4) (2010), 290–297.
- [13] M. Z. Spivey, Combinatorial sums and finite differences. *Discrete Mathematics*, **307** (24) (2007), 3130–3146.
- [14] M.Z. Spivey, *The Art of Proving Binomial Identities* (1st ed.), Chapman and Hall/CRC, 2019.
- [15] R. P. Stanley, *Enumerative Combinatorics: Volume 1*, Cambridge Studies in Advanced Mathematics.
- [16] K. Yordzhev and L. Totina, Matrix representation of a solution of a combinatorial problem of the group theory, <https://doi.org/10.48550/arXiv.1201.3179>
- [17] D. R. Woodall, Finite sums, matrices and induction, *The Mathematical Gazette*, **65** (432) (1981), 92–103.