A Novel Hypermatrix Product and its Application to Multilinear Mappings

S. S. M. A. C. Senevirathne[1,2]*, A. M. C. U. M. Athapattu[3], K. M. N. M. Chathuranga[3]

[1]*Postgraduate Institute of Science, University of Peradeniya, Peradeniya, Sri Lanka*
[2]*Department of Mathematics and Statistics, Faculty of Humanities and Sciences, Sri Lanka Institute of Information Technology (SLIIT), Malabe, Sri Lanka*
[3]*Department of Mathematics, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka*

Corresponding author*: aruni.se@sliit.lk

Abstract

Matrix theory provides a well-established algebraic framework for working with linear maps, in which matrix multiplication replaces the composition of linear transformations. However, there is no canonical multiplication rule for hypermatrices that leads to multilinear maps, partly because multilinear maps are not closed under composition. To address this gap, this research introduces a novel (restricted) hypermatrix multiplication based on the Frobenius inner product. We start by showing that every multilinear map $f: V_1 \times V_2 \times \ldots \times V_n \to V_0$ gives a hypermatrix representation $\mathcal{A}$ and defining a contraction operation, which computes $f(v_1, v_2, \ldots, v_n)$ through Frobenius inner products between $\mathcal{A}$ and matrices derived from input vectors. This operation allows for the efficient computation of the hypermatrix of an arbitrary multilinear map. This work provides constructive proofs and detailed numerical examples.

Keywords: Frobenius Inner Product, Hypermatrices, Linear Transformation; Multilinear Algebra, Multilinear Maps

Introduction

The theory of matrices originated as a tool for solving systems of linear equations. While early appearances of matrix-like methods can be traced back to Chinese mathematics, such as in the *Nine Chapters on the Mathematical Art* (3rd century BCE to 2nd century CE), the formal development of matrix theory is relatively recent. Determinants appeared first in the 17th century in the work of Leibniz and later Cramer (Cajori, 1999). The terminology and conceptual framework of matrices followed in the 19th century, with James Joseph Sylvester (Sylvester, 1850) introducing the term *matrix* in 1850 and Arthur Cayley (Cayley, 1858) developing the basic operations of matrix algebra, including matrix multiplication, in his 1858 *Memoir on the Theory of Matrices* Cayley's insight that compositions of linear transformations correspond to products of matrices laid the foundation for a general algebraic theory. The link between matrices and linear transformations on finite-dimensional vector spaces led to the modern framework of linear algebra, which now pervades both pure and applied mathematics. Matrix theory has become a mature field, complete with rich connections to group theory, geometry, numerical and functional analysis, and physics, particularly quantum mechanics. The development of efficient matrix decompositions (e.g., LU, QR, SVD) and spectral theory has made matrices indispensable in computational mathematics and data science.
In contrast, the theory of hypermatrices (also called multidimensional arrays or tensors) remains comparatively underdeveloped. While higher-order arrays arise naturally in numerous fields such as

https://doi.org/10.54389/KMZF1496

signal processing, quantum information, numerical multilinear algebra, and data analysis, a satisfactory generalization of matrix algebra to higher dimensions is still lacking. A major challenge in developing a comprehensive hypermatrix theory is the absence of a meaningful notion of multiplication. While the matrix product corresponds to the composition of linear maps, there is no canonical or universally accepted way to define a multiplication operation for hypermatrices that retain essential properties such as associativity, bilinearity, and interpretability as multilinear transformations (Kolda & Bader, 2019). Various ad hoc products, such as the outer product and Kronecker product, have been proposed in the literature, often driven by isolated applications rather than guided by a coherent algebraic theory. The absence of a meaningful algebraic framework has limited both the theoretical understanding and practical utility of hypermatrices. This underappreciation is not due to a lack of importance, but rather due to the mathematical difficulties in generalizing familiar matrix notions in meaningful and nontrivial ways.
This study aims to contribute to this ongoing effort by proposing a novel multiplication operation for hypermatrices. Our goal is to retain certain desirable algebraic properties while also ensuring interpretability and relevance to multilinear transformations.

Preliminaries

Definition (Hypermatrices)
Let K be any field. A function form $A : \{0, \dots, n_0 - 1\} \times \{0, \dots, n_1 - 1\} \times \dots \times \{0, \dots, n_{m-1} - 1\} \to K$ is called a hypermatrix of order m and size $n_0 \times n_1 \dots \times n_{m-1}$ with coefficients from K.

Definition (Multilinear map)
A multilinear map is a function that takes multiple vector inputs and is linear in each variable separately. Formally, let $V_0, V_1, \dots, V_n$ be finite-dimensional vector spaces over the same field K. A function,

$$f: V_1 \times V_2 \times \cdots \times V_n \to V_0$$

is said to be multilinear if, for each $1 \le i \le n$, all $\alpha, \beta \in K$ and all vectors $v_1 \in V_1, \dots, v_n \in V_n$, we have:

$$f(v_1, \dots, \alpha v_i + \beta v_i', \dots, v_n) = \alpha f(v_1, \dots, v_i, \dots, v_n) + \beta f(v_1, \dots, v_i', \dots, v_n),$$

Whenever $v_i, v_i' \in V_i$.

Theorem
Let $V_1, V_2, \dots, V_n$ be finite dimensional vector spaces. Let $dim(V_i) = d_i$ for all $0 \le i \le n$, and let $\left(e_{i,j_i}\right)_{1 \le j_i \le d_i}$ be fixed ordered bases for each V_i , respectively.
Let $f: V_1 \times V_2 \times \dots \times V_n \to V_0$ be a multilinear map. Then, there exist scalars, $\left\{ A^{j_0}_{j_i,\dots,j_n} \,\middle|\, 1 \le j_k \le d_k \right\}$ such that,

$$f(v_1, v_2, \dots, v_n) = \sum_{1 \le j_0 \le d_0} \left(\sum_{1 \le j_1 \le d_1} \cdots \sum_{1 \le j_n \le d_n} A^{j_0}_{j_1,\dots,j_n} v_{1,j_1} \cdots v_{n,j_n} \right) e_{0,j_0}$$

where, $v_i = \sum_{1 \le j_i \le d_i} v_{i,j_1} e_{i,j_i}$ for all $1 \le i \le n$ (Serge Lang.,1986)

Proof:
By Mathematical Induction,

For $n = 1$,
Let $f: V_1 \to V_0$ be a linear map. It is well-known that after choosing an ordered basis, for both domain and codomain, f reduces to a matrix. For basis vector $e_{1,j_1} \in V_1$, we define scalars $A^{j_0}_{j_1}$ satisfying,

$$f\left(e_{1,j_1}\right) = \sum_{1 \le j_0 \le d_0} A^{j_0}_{j_1} e_{0,j_0}$$

For any $v_1 = \sum_{1\le j_1\le d_1} v_{1,j_1}\, e_{1,j_1}$ linearity gives,

$$f(v_1) = f\left(\sum_{1\le j_1\le d_1} v_{1,j_1}\, e_{1,j_1}\right) = \sum_{1\le j_1\le d_1} v_{1,j_1}\, f\left(e_{1,j_1}\right) = \sum_{1\le j_1\le d_1} v_{1,j_1}\left(\sum_{1\le j_0\le d_0} A_{j_1}^{j_0}\, e_{0,j_0}\right)$$
$$= \sum_{1\le j_0\le d_0}\left(\sum_{1\le j_1\le d_1} A_{j_1}^{j_0}\, v_{1,j_1}\right) e_{0\,j_0}$$

This proves the theorem for $n = 1$.

Consider the first non-trivial case: $n = 2$,
Let $f: V_1 \times V_2 \to V_0$ be bilinear.
Then the hypermatrix is a 3-dimensional array $(d_0 \times d_1 \times d_2)$ with entries $A_{j_1,j_2}^{j_0}$ defined by,

$$f\left(e_{1,j_1}, e_{2,j_2}\right) = \sum_{1\le j_0\le d_0} A_{j_1,j_2}^{j_0}\, e_{0,j_0}$$

$$f(v_1, v_2) = \sum_{1\le j_0\le d_0}\left(\sum_{1\le j_1\le d_1}\sum_{1\le j_2\le d_2} A_{j_1,j_2}^{j_0}\, v_{1,j_1} v_{2,j_2}\right) e_{0,j_0}$$

This generalizes matrix multiplication to two input vectors.

Then, assume that the theorem is true for multilinear maps with $n-1$ arguments.
i.e , For any multilinear map $g: V_1 \times \dots \times V_{n-1} \to V_0$ there exists a hypermatrix $\mathcal{A}$ with size $d_0 \times d_1 \times \dots \times d_{n-1}$ such that,

$A_{j_1,\dots,j_{n-1}}^{j_0}$ such that,

$$g(w_1, \dots, w_{n-1}) = \sum_{1\le j_0\le d_0}\left(\sum_{1\le j_1\le d_1}\cdots\sum_{1\le j_{n-1}\le d_{n-1}} A_{j_1,\dots,j_{n-1}}^{j_0}\, w_{1,j_1} \dots w_{n-1,j_{n-1}}\right) e_{0,j_0}$$

For n,
Fix $v_n \in V_n$ and define a new map $g_{v_n}: V_1 \times \dots \times V_{n-1} \to V_0$ by,

$$g_{v_n}(v_1, \dots, v_{n-1}) = f(v_1, \dots, v_{n-1}, v_n)$$

Since f is multilinear, g_{v_n} is also linear on its domain.

By the induction hypothesis g_{v_n} has a hypermatrix representation $\mathcal{B}^{v_n}$ of size $d_0 \times d_1 \times \dots \times d_{n-1}$.

$$g_{v_n}(v_1, \dots, v_{n-1}) = \sum_{1\le j_0\le d_0}\left(\sum_{1\le j_1\le d_1}\cdots\sum_{1\le j_n\le d_{n-1}} B_{j_1,\dots,j_{n-1}}^{j_0}\, v_{1,j_1}, \dots, v_{n-1,j_{n-1}}\right) e_{0,j_0}$$

Because g_{v_n} is linear in v_n (by multilinearity of f) the entries $\mathcal{B}_{j_1,\dots,j_{n-1}}^{j_0}$ can be written as,

$$\mathcal{B}_{j_1,\dots,j_{n-1}}^{j_0} = \sum_{1\le j_n\le\le n} A_{j_1,\dots,j_n}^{j_0}\, v_{n,j_n}$$

where $A_{j_1,\dots,j_n}^{j_0}$ are constants.

$$V_n = \sum_{1\le j_n\le d_n} v_{n,j_n}\, e_{n,j_n}$$

Combine terms,

$$f(v_1, \dots, v_n) = \sum_{1\le j_0\le d_0}\left(\sum_{1\le j_1\le d_1}\cdots\sum_{1\le j_{n-1}\le d_{n-1}}\left(\sum_{1\le j_n\le d_n} A_{j_1,\dots,j_n}^{j_0}\, v_{n,j_n}\right) v_{1,j_1}, \dots, v_{n-1,j_{n-1}}\right) e_{0,j_0}$$

Rearrange the sums to get the final hypermatrix,

$$f(v_1, \dots, v_n) = \sum_{1 \le j_0 \le} \left(\sum_{1 \le j_1 \le d_1} \cdots \sum_{1 \le j_n \le d_n} A^{j_0}_{j_1,\dots,j_n} v_{1,j_1}, \dots, v_{n,j_n} \right) e_{0,j_0}$$

Results and Discussion

It is well known that for any bilinear map of the form $f: \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ there is a unique matrix $M = (A_{i,j})_{m \times n}$ such that,

$$f(x, y) = x^T M y = \sum_{1 \le i \le m} \sum_{1 \le j \le n} A_{i,j}\, x_i y_j$$

Observe that with the previous theorem, we can easily derive this result as a corollary. Moreover, it hints at the fact that we can somehow organize the scalars $\left\{ A^{j_0}_{j_1,\dots,j_n} \mid 1 \le j_k \le d_k \right\}$ in the theorem as a hypermatrix.

To illustrate this with an example, let's consider the multilinear map $f: \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^3 \to \mathbb{R}^3$

$$f(u, v, w) = \begin{bmatrix} u_1 v_1 w_1 + u_2 v_2 w_3 \\ 2u_1 v_2 w_2 + u_2 v_1 w_1 \\ u_1 v_1 w_2 + 2u_2 v_2 w_2 \end{bmatrix}$$

Assume standard bases for $\mathbb{R}^2$ and $\mathbb{R}^3$ and compute the hypermatrix entries.

The hypermatrix $\mathcal{A}$ is order $3 \times 2 \times 2 \times 3$, each entry $A^{j_0}_{j_1,j_2,j_3}$ is the coefficient of e_{0,j_0} in $f(e_{1,j_1}, e_{2,j_2}, e_{3,j_3})$.

For $e_{1,1}, e_{2,1}, e_{3,1}$,

$$f\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

All the outputs in $\mathbb{R}^3$ can be calculated the same way.

Then organize $\mathcal{A}$ as a $3 \times 2 \times 2 \times 3$. Each 'slice' corresponds to an output component.

Output $dim1(e_{0,1})$

$$\left[\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}_{(2\times3)}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}_{(2\times3)} \right]$$

Similarly other slices can be calculated.

Hence the Hypermatrix is,

$$H_1 = \left[\left[\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right], \left[\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right], \left[\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} \right] \right]$$

By Frobenius inner product let's consider the slice wise multiplication with the input hypermatrix,

$$\left[\begin{bmatrix} u_1 v_1 w_1 & u_1 v_1 w_2 & u_1 v_1 w_3 \\ u_1 v_2 w_1 & u_1 v_2 w_2 & u_1 v_2 w_3 \end{bmatrix}_{(2\times3)}, \begin{bmatrix} u_2 v_1 w_1 & u_2 v_1 w_1 & u_2 v_1 w_3 \\ u_2 v_2 w_1 & u_2 v_2 w_2 & u_2 v_2 w_3 \end{bmatrix}_{(2\times3)} \right]_{(2\times2\times3)} = u_1 v_1 w_1 + u_2 v_2 w_3$$

Similarly, we can obtain the other components for slices, and therefore we are able to recover the same multilinear map again.

Then let's consider the fixed standard input bases for $\mathbb{R}^2$ and $\mathbb{R}^3$ and consider a non-standard basis for the codomain $\mathbb{R}^3$ as,

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

The corresponding hypermatrix H_2 can be written as follows with respect to the non-standard basis in the codomain.

$$H_2 = \left[\left[\begin{bmatrix}1 & -1 & 0\\ 0 & 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0 & 0\\ 0 & -2 & 1\end{bmatrix}\right], \left[\begin{bmatrix}0 & -1 & 0\\ 0 & 2 & 0\end{bmatrix}, \begin{bmatrix}1 & 0 & 0\\ 0 & -2 & 0\end{bmatrix}\right], \left[\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0 & 0\\ 0 & 2 & 0\end{bmatrix}\right]\right]$$

The following commutative diagram illustrates the relationship between hypermatrix representation of a fixed multilinear map $f: \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^3 \to \mathbb{R}^3$ under different basis choices for the codomain. Here the transformation matrix M, whose columns are the basis vectors of $\mathcal{B}$, converts coordinates from $\mathcal{B}$ to the standard basis. Also, the relationship $H_1 = M \circ H_2$ holds, where $'\circ'$ denotes a scalar-type multiplication (mode-1 product) between the matrix and the hypermatrix. This shows two fundamental operations in our work such as the Frobenius inner product produces hypermatrix with input vectors to evaluate multilinear maps and scalar type multiplication facilitates basis transformations by matrix-hypermatrix composition.

$$\begin{tikzcd}
\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^3 \arrow[rr, "H_2"] \arrow[dr, "H_1"'] & & \mathbb{R}^3 \arrow[dl, "M"] \\
& \mathbb{R}^3 &
\end{tikzcd}$$

Figure 1: *Commutative diagram for hypermatrix basis change*

$$M = \begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}$$

$H_1 = M \circ H_2$

Slice 1;

$$H_1^{(1)} = \left\{1 \circ \left[\begin{bmatrix}1 & -1 & 0\\ 0 & 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0 & 0\\ 0 & -2 & 1\end{bmatrix}\right] + 0 \circ \left[\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\right] + 1 \circ \left[\begin{bmatrix}1 & 1 & 0\\ 0 & 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0 & 0\\ 0 & 2 & 0\end{bmatrix}\right]\right\}$$

$$= \left[\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 1\end{bmatrix}\right]$$

Similarly, the other two slices can be calculated.

Conclusion and Future Work

This work demonstrates that every multilinear map equivalent to a hypermatrix representation, generalizing the concept of matrix multiplication to higher dimensions while preserving linearity in each input. Instead of studying the hypermatrix directly, as there's a correspondence we can discuss multilinear maps. Future work will investigate properties such as invertibility, determinants, and eigenvalues in the context of hypermatrices.

References

Cajori, F. (1999). *A history of mathematics, 303*. American Mathematical Soc.

Cayley, A. (1857, December). A memoir on the theory of matrices. *Proceedings of the Royal Society of London Series I*, *9*, 100-101.

da Silva, J. D., & Machado, A. (2013). *Multilinear algebra. Handbook of Linear Algebra*.

Kolda, T. G., & Bader, B. W. (2009). Tensor decompositions and applications. *SIAM Review, 51*(3), 455-500.

Lang, Serge. *Introduction to Linear Algebra*. 2nd ed. Springer, 1986.

Sylvester, J. J. (1850). XLVII. Additions to the articles in the September number of this journal, “On a new class of theorems,” and on Pascal's theorem. *The London, Edinburgh, and Dublin Philosophical Magazine*