EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
Vol. 7, No. 4, 2014, 405-411
ISSN 1307-5543 – www.ejpam.com

Commutative Law for the Multiplication of Matrices as Viewed in Terms of Hankel's Principle

Yukio Kobayashi

Department of Information Systems Science, Faculty of Engineering, Soka University, Tokyo, Japan

---

Abstract. Many rules of arithmetic for real numbers also hold for matrices, but a few do not. The commutative law for the multiplication of matrices, however, can be also considered as an extension of the law for real numbers. The transpose of a matrix conserves "the principle of the permanence of form and its transition" for the commutative law for multiplication.

2010 Mathematics Subject Classifications: 15A30, 97H60

Key Words and Phrases: Hankel's principle, the principle of the permanence of form and its transition, homothety, matrix, commutative law for multiplication, composite mapping

---

1. Introduction

Any rule of numerical operations should be extended in accordance with Hankel's principle, that is, "the principle of the permanence of form and its transition" [2, 3]. The standard algebraic properties of addition and multiplication are commutativity, associativity, and distributivity. The definitions of addition and multiplication for vectors and matrices should be extended in such a way as to conserve the standard algebraic properties of these numerical operations. The commutative law for multiplication, $ab = ba$, holds for any real numbers a and b. However, $AB = BA$ need not hold for matrices A and B [1]. It seems that the commutative law for multiplication does not follow "the principle of the permanence of form and its transition". The purpose of the present article is to show another view that the commutative law for multiplication also follows this principle through the transpose of a matrix.

2. Composite Mapping as an Extension of a Concept of Proportion

The theory of quantity originated from the problem of proportion. As one variable x doubles, triples, ..., another variable y doubles, triples, ..., respectively. A proportional relation is expressed as a linear equation $y = ax$, where a is a constant. In view of "the principle of

Email address: `koba@t.soka.ac.jp`

the permanence of form and its transition", this relationship is a special case of the theorem that any linear mapping can be represented as a matrix multiplication.

$$\underset{\in\mathbf{R}}{y} = a \underset{\in\mathbf{R}}{x} \longrightarrow \underset{\in\mathbf{R}}{y} = \mathbf{a}' \underset{\in\mathbf{R^n}}{\mathbf{x}} \longrightarrow \underset{\in\mathbf{R^n}}{\mathbf{y}} = A \underset{\in\mathbf{R^n}}{\mathbf{x}}$$

Here $\mathbf{a}'$ is a n-component row vector, $\mathbf{x}$ and $\mathbf{y}$ are n-component column vectors, and A is an $n \times n$ matrix.

The composition of two or more mappings involves taking the output of one or more mappings as the input of other mappings. The mappings $f : X \to Y$ and $g : Y \to Z$ can be composed by first applying f to an argument x to obtain $y = f(x)$ and then applying g to y to obtain $z = g(y)$. The extension of the composition of maps can be expressed in matrix form.

$$\underset{\in\mathbf{R}}{y} = ba \underset{\in\mathbf{R}}{x} \longrightarrow \underset{\in\mathbf{R^n}}{\mathbf{y}} = \mathbf{ba}' \underset{\in\mathbf{R^n}}{\mathbf{x}} \longrightarrow \underset{\in\mathbf{R^n}}{\mathbf{y}} = BA \underset{\in\mathbf{R^n}}{\mathbf{x}}$$

Here b is a constant, $\mathbf{b}$ is an n-component column vector, and B is an $n \times n$ matrix. Thus, the above matrix multiplication is a representation of the composite mapping.

Remark 1. $\mathbf{a}'\mathbf{x}$ *is a scalar, and thus it might seem that* $\mathbf{b}'(\mathbf{a}'\mathbf{x})$ *is a* $1 \times n$ *matrix. This view, however, is not correct, because the associative law of multiplication does not hold for* $\mathbf{b}'\mathbf{a}'\mathbf{x}$. *The multiplication of the* $1 \times n$ *matrices,* $\mathbf{b}'$ *and* $\mathbf{a}'$, *is not defined and thus we cannot calculate* $(\mathbf{b}'\mathbf{a}')\mathbf{x}$. *The disagreement of* $\mathbf{b}'(\mathbf{a}'\mathbf{x})$ *and* $(\mathbf{b}'\mathbf{a}')\mathbf{x}$ *is due to the multiplication of the* $1 \times n$ *matrix* $\mathbf{b}'$ *and the* 1×1 *matrix* $\mathbf{a}'\mathbf{x}$ *in the order violating the rule of matrix multiplication. A detailed discussion is given later (see Section 3).*

The simplest form of a matrix is a 1×1 matrix. If $A = (a)$ and $B = (b)$, the following are true.

$$\begin{aligned}
(AB)^T &= ((a)(b))^T \\
&= (ab)^T \\
&= (ab) \\
&= (a)(b), \\
B^T A^T &= (b)^T (a)^T \\
&= (b)(a),
\end{aligned}$$

where the transposes of A, B, and AB are denoted by A^T, B^T, and $(AB)^T$, respectively. For 1×1 matrices, $(AB)^T = B^T A^T$ can be written as $(a)(b) = (b)(a)$, which can be regarded as $ab = ba$. Therefore, we can consider $(AB)^T = B^T A^T$ as an extension of the commutative law for the multiplication of real numbers, $ab = ba$. In other words, $ab = ba$ is a special case of $(AB)^T = B^T A^T$. Taking this view, "the principle of the permanence of form and its transition" also holds for the commutative law for matrix multiplication.

3. Scalar Multiplication of a Matrix

The transpose of a matrix plays an essential role in maintaining "the principle of the permanence of form and its transition" for matrix multiplication. As shown in the previous section,

$\mathbf{ba'x}$ is valid, whereas $\mathbf{b'a'x}$ is not. The essence of the reason is that scalar multiplication can be treated as only an abbreviation to indicate multiplication by a scalar matrix, which is a diagonal matrix whose diagonal elements all contain the same scalar. Here, we express $\mathbf{a'x}$ as a scalar λ. For simplicity, let us consider the case $n = 3$. Then, $\mathbf{b}'\lambda = (b_1 \ \ b_2 \ \ b_3)\lambda = (b_1\lambda \ \ b_2\lambda \ \ b_3\lambda)$ is a convenient operation, but deviates from the rule of matrix multiplication. The proper operation is

$$(b_1 \ \ b_2 \ \ b_3)\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} = (b_1\lambda \ \ b_2\lambda \ \ b_3\lambda),$$

because the multiplication of the 1×3 matrix $\mathbf{b}'$ and the 1×1 matrix λ is not defined. A scalar λ in $\mathbf{b}'\lambda$ is an abbreviation of a scalar matrix Λ, where

$$\Lambda = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix},$$

and thus scalar multiplication implies a mapping referred to as homothety of ratio λ. If we apply the rule of matrix multiplication properly, we are easily convinced that $\mathbf{b'a'x}$ is not valid. The transpose of $\mathbf{b}'\lambda$ is $\lambda\mathbf{b}$, which is an abbreviation of

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

The commutative law

$$(b_1 \ \ b_2 \ \ b_3)\lambda = \lambda(b_1 \ \ b_2 \ \ b_3)$$

indicates that

$$\begin{aligned}(b_1\lambda \ \ b_2\lambda \ \ b_3\lambda) =& (b_1 \ \ b_2 \ \ b_3)\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} = \left[\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\right]^T \\ =& (\lambda b_1 \ \ \lambda b_2 \ \ \lambda b_3).\end{aligned}$$

Similarly, the commutative law

$$\lambda\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\lambda$$

indicates that

$$\begin{pmatrix} \lambda b_1 \\ \lambda b_2 \\ \lambda b_3 \end{pmatrix} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \left[(b_1 \ \ b_2 \ \ b_3)\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}\right]^T = \begin{pmatrix} b_1\lambda \\ b_2\lambda \\ b_3\lambda \end{pmatrix}.$$

These forms show an extension of the commutative law for the multiplication of real numbers, $\lambda b = b\lambda$. If linear transformation is homothety, $(b_1\ \ b_2\ \ b_3)\lambda = \lambda(b_1\ \ b_2\ \ b_3)$ and $\lambda\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\lambda$ indicates the left-handed vector space can be equated with the right-handed vector space. The right multiplication of a row vector with a scalar, $\mathbf{b}'\lambda$, and the left multiplication of a column vector with a scalar, $\lambda\mathbf{b}$, are not a rule of matrix arithmetic but the abbreviations of $\mathbf{b}'\Lambda$ and $\Lambda\mathbf{b}$, respectively, and thus the vectors obtained by multiplying each entry of $\mathbf{b}'$ and $\mathbf{b}$ by λ are the definition of scalar multiples.

4. Related Remarks

At the high school and undergraduate level, the scalar multiplication of a column vector is expressed as

$$\lambda\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix},$$

where λ is a scalar. As shown in the previous section, this form is not proper from the viewpoint of matrix multiplication. In standard textbooks on linear algebra (for example, [1]), however, this form is used as the first step of the procedures for diagonalizing a matrix and that for deriving the standard matrix for a rotation operator. These procedures are not necessarily easy for some students for the following reason.

Diagonalization of a Matrix

For an $n \times n$ diagonalizable matrix A, there is an invertible matrix

$$U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{pmatrix},$$

where $\begin{pmatrix} u_{11} \\ u_{21} \\ \vdots \\ u_{n1} \end{pmatrix}, \begin{pmatrix} u_{12} \\ u_{22} \\ \vdots \\ u_{n2} \end{pmatrix}, \ldots, \begin{pmatrix} u_{1n} \\ u_{2n} \\ \vdots \\ u_{nn} \end{pmatrix}$ are the eigenvectors of A, such that matrix U diagonalizes A, that is, $U^{-1}AU = \Lambda$, where

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

and $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A. This raises a question: It follows from the formula $U^{-1}AU = \Lambda$ that $AU = U\Lambda$. We have $A\mathbf{u}_1 = \lambda_1\mathbf{u}_1$, $A\mathbf{u}_2 = \lambda_2\mathbf{u}_2$, ..., $A\mathbf{u}_n = \lambda_n\mathbf{u}_n$, where $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ are the eigenvectors of A corresponding to the eigenvalues of A, $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively. The question is, why do we consider $U\Lambda$ instead of ΛU as equivalent to the forms of $\lambda_1\mathbf{u}_1, \lambda_2\mathbf{u}_2, \ldots, \lambda_n\mathbf{u}_n$? We can avoid this question by following the rule of matrix multiplication. The forms $\lambda_1\mathbf{u}_1, \lambda_2\mathbf{u}_2, \ldots, \lambda_n\mathbf{u}_n$ deviate from this rule because the multiplication of a 1×1 matrix and an $n \times 1$ matrix is not defined. Therefore, if we consider $\mathbf{u}_1\lambda_1, \mathbf{u}_2\lambda_2, \ldots, \mathbf{u}_n\lambda_n$ instead of the above forms, a combined form $U\Lambda$, or

$$\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

can easily be obtained.

Some instructors do not regard eigenvalues as matrices and explain that λ_k is just taken from left to right in actually operating with the commutativity as it is a scalar. This interpretation is ambiguous in meaning, although we can consider that the column vectors of the product $U\Lambda$ are $\lambda_1\mathbf{u}_1, \lambda_2\mathbf{u}_2, \ldots, \lambda_n\mathbf{u}_n$. From a pedagogical standpoint, it is not necessarily easy for some students to understand the process of the construction of the combined form $U\Lambda$ from $\lambda_1\mathbf{u}_1, \lambda_2\mathbf{u}_2, \ldots, \lambda_n\mathbf{u}_n$. Properly, the right multiplication of a column vector $\mathbf{u}_k$ with a scalar λ_k, $\mathbf{u}_k\lambda_k$, implies a mapping referred to as homothety of ratio λ_k, and thus eigenvalues can be regarded as 1×1 matrices for representing linear maps. The right multiplication of a column vector with a scalar is compatible with the definition of matrix multiplication, because the number of columns of the column matrix (or the column vector) is the same as the number of row of the 1×1 matrix (or the scalar). The column vectors of the product $U\Lambda$ are $\mathbf{u}_1\lambda_1, \mathbf{u}_2\lambda_2, \ldots, \mathbf{u}_n\lambda_n$. We can also rewrite $A\mathbf{u}_k = \mathbf{u}_k\lambda_k$ as $A\mathbf{u}_k = \Lambda_k\mathbf{u}_k$, where Λ_k is a scalar matrix whose diagonal entries are equal to λ_k, because

$$\begin{pmatrix} u_{1k} \\ u_{2k} \\ \vdots \\ u_{nk} \end{pmatrix} \lambda_k = \begin{pmatrix} \lambda_k & 0 & \cdots & 0 \\ 0 & \lambda_k & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{pmatrix} \begin{pmatrix} u_{1k} \\ u_{2k} \\ \vdots \\ u_{nk} \end{pmatrix}.$$

See Section 3 for this. This relationship indicates that homothety can be represented by a scalar matrix. To find the eigenvalues and eigenvectors of A, we rewrite $A\mathbf{u}_k = \mathbf{u}_k\lambda_k$ as $(A - \Lambda_k)\mathbf{u}_k = \mathbf{0}$ and obtain the characteristic equation of A. By transforming $A\mathbf{u}_k = \mathbf{u}_k\lambda_k$ into $(A - \Lambda_k)\mathbf{u}_k = \mathbf{0}$, we solve the problem to obtain the kernel of a linear transformation expressed by the $n \times n$ matrix $A - \Lambda_k$. We take note of the relationship $\mathbf{u}_k\lambda_k = \Lambda_k\mathbf{u}_k$ only in the process of obtaining the characteristic equation of A.

Standard Matrix for Rotation Operator

For simplicity, we consider the rotation operator on $\mathbf{R}^2$. A vector $\mathbf{r}$ can be expressed as

$$\mathbf{r} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} x + \begin{pmatrix} 0 \\ 1 \end{pmatrix} y.$$

Each term on the right-hand side is in the form of a multiplication of a 2×1 matrix and a 1×1 matrix. Application of the rotation operator that rotates each vector counterclockwise through a fixed positive angle θ yields

$$\mathbf{r}' = \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} x + \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} y,$$

which can be easily rewritten in matrix form as

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

If we express $\mathbf{r}$ as

$$\mathbf{r} = \begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

we obtain

$$\mathbf{r}' = \begin{pmatrix} x' \\ y' \end{pmatrix} = x \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} + y \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}.$$

In expressing this equation in matrix form, we must rearrange the order of x, y, and the trigonometric functions. Thus, orders violating the rule of matrix multiplication are inconvenient. Similarly, the simultaneous equations for obtaining the coefficients of a linear combination of column vectors,

$$c_1 \begin{pmatrix} 5 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 7 \\ 3 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix},$$

are

$$\begin{cases} 5c_1 + 7c_2 = 4 \\ 2c_1 + 3c_2 = 6 \end{cases},$$

in which $c_1 5$, $c_2 7$, and so on are commuted into $5c_1$, $7c_2$, and so on. These simultaneous equations are equivalent to that using the transpose of the original column vectors,

$$(5 \ \ 2)c_1 + (7 \ \ 3)c_2 = (4 \ \ 6).$$

If we express the vector equation as

$$\begin{pmatrix} 5 \\ 2 \end{pmatrix} c_1 + \begin{pmatrix} 7 \\ 3 \end{pmatrix} c_2 = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$

following the rule of matrix multiplication, we need not commute the order of multiplication in the simultaneous equations.

5. Conclusion

According to the standard interpretation, the commutative law for multiplication, $AB = BA$, is not valid in matrix arithmetic. Instead, we can interpret that $(AB)^T = B^T A^T$ is a rule of matrix arithmetic. The commutative law for multiplication, $ab = ba$, for any real numbers a and b can be regarded as a spacial case of $(AB)^T = B^T A^T$. The transpose of a matrix conserves "the principle of the permanence of form and its transition" for the commutative law for multiplication. This view indicates that we can unify a rule of matrix arithmetic and that of real number algebra to avoid the exception.

From this point of view, the right multiplication of a column vector with a scalar is the matrix multiplication of a $n \times 1$ matrix and a 1×1 matrix, which is compatible with the definition of matrix multiplication in contrast to the left multiplication of a column vector with a scalar. For example, as shown in Section 4, the process of finding a corresponding diagonal matrix for a diagonalizable matrix becomes clear by considering a scalar multiple the right multiplication of a column vector with a scalar.

If we omit the brackets on a 1×1 matrix, it is impossible to distinguish between the number and the 1×1 matrix whose entry takes the same value as the number. However, it is usually possible to tell which is meant from the context in which the symbol appears [1]. Scalar multiplication implies homothety, and thus the scalar can be regarded as a 1×1 matrix or an abbreviation of a scalar matrix as shown in Section 3.

References

[1] H Anton and C Rorres. *Elementary Linear Algebra Applications Version*. John Wiley & Sons, New York, 1994.

[2] E E Escultura. Creative Mathematics Education. *Creative Education*, **3**, 45-54 (2012).

[3] Y Itagaki. The Concept of Rational Numbers: A study of the View on the Principle of the Permanence of Form and Its Transition. *Journal of Japan Society of Mathematical Education*, **67**, 114-121 (1985) [in Japanese].