APPLICATIONS OF MATRICES MULTIPLICATION TO DETERMINANT AND ROTATIONS FORMULAS IN $\mathbb{R}^n$

ALEX GOLDVARD AND LAVI KARP

Abstract. This note deals with two topics of linear algebra. We give a simple and short proof of the multiplicative property of the determinant and provide a constructive formula for rotations. The derivation of the rotation matrix relies on simple matrix calculations and thus can be presented in an elementary linear algebra course. We also classify all invariant subspaces of equiangular rotations in 4D.

1. Introduction

This article aims to promote several geometric aspects of linear algebra. The geometric motivation often leads to simple proofs in addition to increasing the students’ interest of this subject and providing a solid basis. Students with a confident grasp of these ideas will encounter little difficulties in extending them to more abstract linear spaces. Many geometrical operations can be rephrased in the language of vectors and matrices. This includes projections, reflections and rotations, operations which have numerous applications in engineering, physics, chemistry and economic, and therefore their matrix’s representation should be included in basic linear algebra course.

There are basically two attitudes in teaching linear algebra. The abstract one which deals with formal definitions of vector spaces, linear transformations etc. Contrary to the abstract vector spaces, the analytic approach deals mainly with the vector space $\mathbb{R}^n$ and provides the basic concepts and proofs in these spaces.

However, when the the analytic approach deals with the definition of linear transformations it uses the notion of representation of the matrix of the transformation in arbitrary basis and thus it actually goes back to the abstract setting. To clarify this issue, let us consider for example the calculations a rotation $T$ of a vector $\mathbf{x}$ in $\mathbb{R}^3$. In this way one start

2010 Mathematics Subject Classification. Primary 15B10, 97H60; Secondary 15A04, 15A15.

Key words and phrases. Rotation matrix, matrices multiplication, determinant, orthogonal matrices, equiangular rotations.
with choosing an appropriate basis $B$, and calculating the matrix of the given transformation in this basis $[T]_B$. In the second step one has to compute the transformation matrix $P$ from the standard basis to the basis $B$ and its inverse $P^{-1}$. The final step consists of applying the matrix $P^{-1}[T]_BP$ to $x$. This cumbersome machinery is common for both the abstract and the analytic approaches. In addition, it takes an essential effort to teach all the necessary details in order to use this non-practicable formula.

It is amazing why one should use this complicated formula while the Rodrigues’ rotation formula does it efficiently. Although its simplicity, Rodrigues’ formula does not appear in the current linear algebra textbooks. Its proof is elementary, but require non-trivial geometric insight.

In this note we present a new proof of the matrix’s representation of Rodrigues’ formula. The essential point is that we regard the multiplication $Ax$, of a vector $x$ by a matrix $A$, simultaneously as an algebraic operation and geometric transformation (similarly to Lay [1]). This enable us to derive the rotation formula in the three dimensional space. In higher dimensional spaces we first propose a geometric definition of a rotation and after that we derive the formula in a similar manner to the three dimensional spaces. Having calculate the matrix of rotation according to that definition, we show it is identical to the common definition of rotation, that is, an orthogonal matrix with determinant one.

We also consider the multiplication of two matrices $AB$ as multiplications of the columns of $B$ by the matrix $A$. Applying this point of view we provide a simple proof of the multiplicative property of the determinants. The standard proof of this property is often being skipped from the classroom since it is considered as too complicated. The proof which we present here could easily be thought in the beginning of a linear algebra course.

2. Basic facts and notations

We recall the definitions of multiplication of a vector by a matrix and the multiplications of two matrices. Both definitions rely solely on the basic two operations of vectors in $\mathbb{R}^n$, namely, addition and multiplication by a scalar.
We denote $m \times n$ matrix $A$ by $[a_1, a_2, \ldots, a_n]$, where \{a_1, a_2, \ldots, a_n\} are the columns of $A$. Let $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ be a vector in $\mathbb{R}^n$, then

(1) \[ A\mathbf{x} = x_1a_1 + x_2a_2 + \ldots + x_na_n. \]

This means that $A\mathbf{x}$ is simply a linear combination of the columns of the matrix $A$. Conversely, any linear combination of $n$ vectors $\{a_1, a_2, \ldots, a_n\}$ can be written as a matrix multiplication. Note that beside of $A\mathbf{x}$ being a linear combination, we can interpret it as a transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ by corresponding to each $\mathbf{x} \in \mathbb{R}^n$ the vectors $A\mathbf{x} = \mathbf{y} \in \mathbb{R}^m$. Obviously this operation has the linearity property:

(2) \[ A(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha A\mathbf{u} + \beta A\mathbf{v}. \]

Thus any linear transformation can be written as a multiplication of vectors by a matrix and therefore the formal definition the linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$ seems to be superfluous.

Most of the textbooks define a matrix multiplications by the row-column rule. But the original definition of Cayley is by means of a composition of two linear substitutions (see e.g. [4, 6]). This means that if $A$ is a $m \times n$ matrix and $B$ $n \times k$, then the matrix $C = AB$ is defined through the identity $C\mathbf{x} = A(B\mathbf{x})$, where $\mathbf{x} \in \mathbb{R}^k$. From (1) and (2) it immediately follows that

(3) \[ AB = [Ab_1, Ab_2, \ldots, Ab_k], \]

where $\{b_1, b_2, \ldots, b_k\}$ are the columns of the matrix $B$.

Another useful way to multiply matrices is by the column row rule, that is,

(4) \[ AB = a_1b_1^T + \ldots + a_nb_n^T, \]

where $b_1^T, \ldots, b_n^T$ are rows vectors of the transpose $B^T$. This type of multiplication will be used in Section 4.

3. Multiplicative property of determinant

Let $A$ be $n \times n$ matrix with coefficients $[a_{ij}]_{i,j=1}^n$. The determinant of $A$ is defined to be the scalar

(5) \[ \det(A) = \sum_p \sigma(p)a_{1p_1}a_{2p_2} \ldots a_{np_n}, \]
where the sum is taken over the \( n! \) permutations \( p = (p_1, p_2, \ldots, p_n) \) of \((1, 2, \ldots, n)\) and

\[
\sigma(p) = \begin{cases} 
+1 & \text{if } p \text{ is even permutation} \\
-1 & \text{if } p \text{ is odd permutation}
\end{cases}
\]

**Theorem 1.** Suppose \( A \) and \([u, v, \ldots, z]\) are two \( n \times n\) matrices, then

\[
(6) \quad \det([Au, Av, \ldots, Az]) = \det(A) \det([u, v, \ldots, z]).
\]

**Proof** Denote \( A \) by \([a_1, a_2, \ldots, a_n]\) and let

\[
u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \quad z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}.
\]

We are now using formula (1), the linearity of the determinant and the fact that a matrix having two equal columns its determinant is zero. All these result with

\[
\det([Au, Av, \ldots, Az]) = \sum_p \sigma(p) \det([a_{p_1}, a_{p_2}, \ldots, a_{p_n}]).
\]

Since the determinant changes sign when two columns are interchanged, we have

\[
\det([a_{p_1}, a_{p_2}, \ldots, a_{p_n}]) = \sigma(p) \det(A).
\]

Therefore,

\[
\det([Au, Av, \ldots, Az]) = \det(A) \sum_p \sigma(p) u_{p_1} v_{p_2} \ldots z_{p_n}
\]

\[
= \det(A) \det([u, v, \ldots, z]).
\]

\[\square\]

The above Theorem has an important geometric interpretation and we discuss it here in \( \mathbb{R}^2 \). It is well known that if \( u, v \) are two collinear vectors, then \( \det([u, v]) \) is the area of parallelogram spanned by \( \{u, v\} \).

Now if \( \det(A) \neq 0 \), then \( \{Au, Av\} \) span also a parallelogram. Therefore the number \( \det(A) \) is the proportion between the areas of the parallelograms spanned by \( \{u, v\} \) and \( \{Au, Av\} \).

Note that Theorem 1 gives the multiplicative property of the determinant. Indeed, let \( B = [u, v, \ldots, z] \), then \( AB = [Au, Av, \ldots, Az] \) and (6) becomes

\[
\det(AB) = \det(A) \det(B).
\]
4. The derivation of a formula for a rotation matrix in $\mathbb{R}^n$

First of all let us discuss the definition of a rotation in $\mathbb{R}^n$. The common definition of a rotation is by means of an orthogonal matrix with determinant one. Here we provide another definition based on geometric considerations. We are aware that such definition was probably given in the past, but we could note traced it. Its advantage is being practicable and the computation of the rotation’s matrix does not utilize eigenvalues and eigenvectors. We shall then verify the equivalence of the two definitions.

It is rather simple to define a rotation in $\mathbb{R}^2$. A linear transformation $R$ is a rotation if the angle between the vectors $R\mathbf{x}$ and $\mathbf{x}$ is a constant for all $\mathbf{x} \in \mathbb{R}^2$. From this definition follows that

$$R \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix},$$

where the rotation is by an angle $\alpha$ counterclockwise. Therefore

$$(7) \quad R = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} = \cos \alpha I + \sin \alpha J$$

where $I$ is the identity matrix and $J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Formula (7) implies that $||R\mathbf{x||} = ||\mathbf{x||}$ for each $\mathbf{x} \in \mathbb{R}^2$ and that confirms our geometric intuition. Matrices which preserve norm are called orthogonal matrices. Their determinant is $\pm 1$ and rotations are orthogonal matrices with determinant one.

The definition of a rotation in $\mathbb{R}^3$ is slightly more involved. A linear transformation $R$ is called rotation if there exists two dimensional subspace $\Pi$ of $\mathbb{R}^3$ (the plane of the rotation) such that the angle between vectors $R\mathbf{x}$ and $\mathbf{x}$ is a constant for all $\mathbf{x} \in \Pi$, and $R\mathbf{y} = \mathbf{y}$ for each $\mathbf{y}$ orthogonal to $\Pi$ (the axis of the rotation). Euler’s theorem about the rigid motion of the sphere with fixed center justifies this definition.

In order to calculate the matrix $R$ we pick two orthonormal vectors $\mathbf{a}, \mathbf{b} \in \Pi$ and a unit vector $\mathbf{c}$ which is orthogonal to $\Pi$ and such that the triple $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ is right-handed. Applying the rotation $R$ to these vectors results with

$$Ra = \cos \alpha a + \sin \alpha b, \quad Rb = -\sin \alpha a + \cos \alpha b, \quad Rc = c.$$

We write the above equalities in a matrix form $RP = Q$, where $P = [\mathbf{a}, \mathbf{b}, \mathbf{c}]$ and $Q = [Ra, Rb, Rc]$. Since $P$ is an orthogonal matrix, $R =$
$QP^T$ and calculating $QP^T$ by means of (4), we get that

$$R = [Ra, Rb, Rc] \begin{bmatrix} a^T \\ b^T \\ c^T \end{bmatrix} = (Ra)a^T + (Rb)b^T + ( Rc)c^T$$

$$= \cos\alpha \left(aa^T + bb^T\right) + \sin\alpha \left(ba^T - ab^T\right) + cc^T.\quad (8)$$

The skew symmetric matrix $(ba^T - ab^T)$ is the matrix representation of the cross product $c \times x$. To see this note that $(ba^T - ab^T)c = 0$, $(ba^T - ab^T)a = b$ and $(ba^T - ab^T)b = -a$. Hence

$$Rx = \cos\alpha x + (1 - \cos\alpha)cc^Tx + \sin\alpha(c \times x),$$

which is the known Rodrigues’ formula.

Formula (8) resembles to a large extent the two dimensional formula (7). The matrix $(aa^T + bb^T)$ is the projection on the plane $\Pi$, $cc^T$ is the projection on the line orthogonal to $\Pi$ and

$$(ba^T - ab^T)^2 = -(aa^T + bb^T).\quad (9)$$

Since the rotation is actually in the plane $\Pi$, we see that $(aa^T + bb^T)$ corresponds $I$ and $(ba^T - ab^T)$ corresponds $J$ in formula (7). In both formulas, $\cos\alpha$ is the coefficient of a symmetric matrix and $\sin\alpha$ is the coefficient of an anti-symmetric matrix.

It is easy to check that the matrix $R(\alpha) := R$ in (8) is an orthogonal matrix with determinate one. Indeed, relation (9) and the orthogonality of $\{a, b, c\}$ imply that

$$R(\alpha)R^T(\alpha) = \left(\cos^2\alpha + \sin^2\alpha\right) (aa^T + bb^T) + cc^T = I$$

and hence $\det(R(\alpha)) = \pm 1$. Letting $\lim_{\alpha \to 0} R(\alpha) = I$ and using the continuity of the determinants, we see that $\det(R(\alpha)) = 1$.

We turn now to rotations in $\mathbb{R}^4$. It turns out that rotations in $\mathbb{R}^4$ can be defined in a similar way to rotations in $\mathbb{R}^2$ and $\mathbb{R}^3$. We say that a linear transformation $R$ is a rotation if there exists two dimensional subspace $\Pi$ of $\mathbb{R}^4$ such that the angle between vectors $Rx$ and $x$ is a constant for all $x \in \Pi$, and the angle between vectors $Ry$ and $y$ is a constant for all $y \in \Pi^\perp$, the orthogonal complement of $\Pi$.

The calculation are done in a similar manner as we did in $\mathbb{R}^3$. Pick $a, b \in \Pi$ and $c, d \in \Pi^\perp$ such that the set $\{a, b, c, d\}$ is an orthonormal basis. Let $R = R(\alpha, \beta)$ be the rotation matrix with rotation’s angles $\alpha$ in the plane $\Pi$ and $\beta$ in the orthogonal complement $\Pi^\perp$. Then

$$Ra = \cos\alpha a + \sin\alpha b, \quad Rb = -\sin\alpha a + \cos\alpha b,$$

$$Rc = \cos\beta c + \sin\beta d, \quad Rd = -\sin\beta c + \cos\beta d.$$
Set $P = [a, b, c, d]$ and $Q = [Ra, Rb, Rc, Rd]$, since the matrix $P$ is orthogonal, $R = QP^T$ and hence

$$
R = (Ra)a^T + (Rb)b^T + (Rc)c^T + (Rd)d^T
$$

(10)

$$
= \cos \alpha (aa^T + bb^T) + \sin \alpha (ba^T - ab^T) \\
+ \cos \beta (cc^T + dd^T) + \sin \beta (dc^T - cd^T).
$$

We can now easily distinguish between two types of 4D-rotations. If $\beta = 0$, then the rotation is simple, that is, $Ry = y$ for all $y \in \Pi^\perp$. Otherwise, both planes $\Pi$ and $\Pi^\perp$ rotate simultaneously and this type is called a double rotation.

If one doubts whether the matrix $R$ in (10) is an orthogonal matrix with determinant one, then the following simple calculation will convince him. Since

$$
R(\alpha, \beta)R^T(\alpha, \beta) = (\cos^2 \alpha + \sin^2 \alpha) (aa^T + bb^T) + (\cos^2 \beta + \sin^2 \beta) (cc^T + dd^T)
$$

(11)

$$
= I,
$$

$R(\alpha, \beta)$ is an orthogonal matrix and by letting $\alpha$ and $\beta$ go to zero, we get that its determinant is one.

We are now in position to extent the geometric definition of rotations to $\mathbb{R}^n$ for arbitrary positive integer $n$. For $n = 2p$ we say that a linear transformation $R$ is a rotation if there exist $p$ mutual orthogonal planes $\Pi_k$ such that the angle between the vectors $Rx$ and $x$ is a constant for all $x \in \Pi_k$, $k = 1, \ldots, p$. For $n = 2p + 1$ we require that there are $p$ mutual orthogonal planes $\Pi_k$ and in addition a line $L$ is orthogonal to $\Pi_k$, $k = 1, \ldots, p$ such that $R$ behaves the same as in the even on the planes $\Pi_k$ and $Ry = y$ for all $y \in L$. The extension of formulas (8) and (10) to arbitrary dimension is obvious. In $\mathbb{R}^{2p}$ there is an orthonormal basis $\{(a_1, b_1), \ldots, (a_n, b_n)\}$ such that

(11)

$$
R = \sum_{k=1}^{p} \cos \alpha_k \left( a_k a_k^T + b_k b_k^T \right) + \sin \alpha_k \left( b_k a_k^T - a_k b_k^T \right).
$$

In odd dimension $2p + 1$ there is an orthonormal basis $\{(a_1, b_1), \ldots, (a_n, b_n), c\}$ such that

(12)

$$
R = \sum_{k=1}^{p} \cos \alpha_k \left( a_k a_k^T + b_k b_k^T \right) + \sin \alpha_k \left( b_k a_k^T - a_k b_k^T \right) + cc^T.
$$

Similarly to the rotation formulas (8) and (10) one can check that (11) and (12) are orthogonal matrices with determinate one.
Formulas (11) and (12) were derived by [5] but in a different way. The advantage of the derivation given here is being constructive an addition to being appropriate for an elementary linear algebra course. Formulas (11) and (12) can be written in a vectors’ form

\[ R\mathbf{x} = \sum_{k=1}^{p} \cos \alpha_k \mathbf{y}_k + \sin \alpha_k \mathbf{z}_k \]

and

\[ R\mathbf{x} = \sum_{k=1}^{p} \cos \alpha_k \mathbf{y}_k + \sin \alpha_k \mathbf{z}_k + (\mathbf{c}^T \mathbf{x}) \mathbf{c}, \]

where \( \mathbf{y}_k \) is the projection of vector \( \mathbf{x} \) on the plane \( \Pi_k \) and \( \mathbf{z}_k \) is the rotation of \( \mathbf{y}_k \) by an angle \( \pi/2 \) in the plane \( \Pi_k \).

4.1. Invariant subspaces of equiangular subspaces of rotations in \( \mathbb{R}^4 \). A rotation \( R \) in \( \mathbb{R}^4 \) is called equiangular rotations or isoclinic rotations if the planes \( \Pi \) and its orthogonal complement \( \Pi^\perp \) rotate with the same angle (see e.g. [2, 3]). When \( \alpha \neq \beta \), then the planes \( \Pi \) and \( \Pi^\perp \) are the only invariant subspaces under the rotation \( R \). However, when \( \alpha = \beta \), then there are infinitely many two dimensional invariant planes (see e.g. [5]). We shall see here that this interesting phenomenon is a simple consequence of the formula (10) and we shall also classify all the invariant planes.

To see this we note that when \( \alpha = \beta \), then (10) becomes

\[ R = \cos \alpha I + \sin \alpha J, \]

where \( I \) is the identity on \( \mathbb{R}^4 \) and \( J = (ba^T - ab^T + dc^T - cd^T) \). Hence a subspace \( U \) of \( \mathbb{R}^4 \) is an invariant subspace of \( R \) if and only if it is invariant subspace of the matrix \( J \).

Now \( J \) is a skew-symmetric matrix satisfying \( J^2 = -I \). Therefore it has no real eigenvalues and this implies that any non-trivial invariant subspace must has dimension two. Since \( J^2 = -I \), \( \text{span}\{\mathbf{u}, J\mathbf{u}\} \) is an invariant subspace for any \( \mathbf{u} \in \mathbb{R}^4 \). On the other hand, if \( U \) is a non-trivial invariant subspace of \( R \), then \( J\mathbf{u} \in U \) for any \( \mathbf{u} \in U \). Hence \( U \) must be spanned by these vectors. Thus we have obtained a complete classification of the invariant subspaces of equiangular rotations which is independent of the rotation angle \( \alpha \).

It follows from formulas (11) and (12) that if all the angles \( \alpha_k \) are equal, then there are infinitely many invariant subspaces and each one of them is spanned by a vector \( \mathbf{u} \) and \( \sum_{k=1}^{p} (b_k\mathbf{a}_k^T - a_k\mathbf{b}_k^T) \mathbf{u} \).
5. Concluding Remarks

Many vector-space textbooks use the entry-by-entry definition $c_{ij} = \sum a_{ik}b_{kj}$ for the matrices multiplications. The operation of multiplication of a vector $\mathbf{x}$ by a matrix $A$ in accordance (1) bears in itself both geometric and algebraic properties. Therefore a decent understanding of it should be prior to the formal definition of matrices multiplication. After that the matrices’ multiplication in Cayles’s spirit follows naturally. The Cayley’s definition (3) and the column-row rule (4) have many advantages. In many cases they makes the computations easier in addition to increases the comprehension. This note emphasizes two aspects of that attitude.

References

1. D.C. Lay, Linear Algebra and Its Applications, 3rd Edition, Addison-Wesley Higher Education Group, 2003.
2. P. Lounesto, Clifford Algebras and Spinors, 2nd Edition, London Mathematical Society, Lecture Notes Series 286, Cambridge Press 2001.
3. H.P. Manning Geometry of Four Dimension, The Macmillan Company, 1914.
4. C.D. Meyer, Matrix Analysis and Applied Linear Algebra, SIAM, 2000.
5. S. A. Schelkunoff, On rotations in ordinary and null spaces, American Journal of Mathematics, 53, No. 1 (1931), 175-185.
6. A. Tucker, The growing importance of linear algebra in undergraduate mathematics, The College Mathematics Journal, 24, No. 1 (1993), 3-9.

E-mail address: goldvard@braude.ac.il

Department of Mathematics, ORT Braude College, P.O. Box 78, 21982 Karmiel, Israel

E-mail address: karp@braude.ac.il

Department of Mathematics, ORT Braude College, P.O. Box 78, 21982 Karmiel, Israel