

THE ROTATION OF A COORDINATE SYSTEM AS A LINEAR TRANSFORMATION¹

Representation of vectors and linear transformation. Let V be a finite dimensional vector space over the field F , and let $\mathcal{X} = (\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^n)$ be a basis of V .² For a vector $\mathbf{x} \in V$ we have $\mathbf{x} = \sum_{i=1}^n c_i \mathbf{x}^i$ for some $c_i \in F$ ($1 \leq i \leq n$). Let \mathbf{c} be the column vector $\mathbf{c} = (c_1, \dots, c_n)^T$, where the superscript T indicates transpose. We will write

$$\mathbf{x} = \sum_{i=1}^n \mathbf{x}^i c_i = \mathcal{X} \mathbf{c},$$

where for the second equation we consider \mathcal{X} a row vector, and then the product on the right-hand side is viewed as a product of a row vector and a column vector. We introduce the notation

$$\mathcal{R}_{\mathcal{X}} \mathbf{x} \stackrel{def}{=} \mathbf{c},$$

and call the column vector \mathbf{c} the representation of the vector \mathbf{x} with respect to the basis \mathcal{X} .

Now, let V and W be vector spaces over F , let $\mathcal{X} = (\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^n)$ be a basis of V , let $\mathcal{Y} = (\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^m)$ be a basis of W , and let $T : V \rightarrow W$ be a linear transformation. For any i with $1 \leq i \leq n$ we have $T\mathbf{x}^i = \sum_{j=1}^m \mathbf{y}^j a_{ji}$ with some $a_{ji} \in F$. For an arbitrary vector $\mathbf{x} \in V$ such that $\mathbf{c} = (c_1, \dots, c_n)^T = \mathcal{R}_{\mathcal{X}} \mathbf{x}$ we have $\mathbf{x} = \sum_{i=1}^n \mathbf{x}^i c_i$ and $T\mathbf{x} = \sum_{j=1}^m \mathbf{y}^j \sum_{i=1}^n a_{ji} c_i$. In matrix form, this equation will be written as $T\mathbf{x} = T\mathcal{X}\mathbf{c} = \mathcal{Y}A\mathbf{c}$ where A is the $n \times m$ matrix with entries a_{ji} in the j th row and the i th column. In the last equation, we may omit the column vector \mathbf{c} on the right-hand side and write $T\mathcal{X} = \mathcal{Y}A$. We will call the matrix A the representation of the linear transformation T with respect to the bases \mathcal{X} and \mathcal{Y} , and write

$$\mathcal{R}_{\mathcal{Y}\mathcal{X}} T \stackrel{def}{=} A.$$

The basic properties of representations of vectors and of linear transformations are discussed in [1].

Addition formulas for trigonometric functions. In using vectors to discuss plane geometry, it is natural to introduce the vector \mathbf{i} , of unit length, pointing in the direction of the positive x axis, and the vector \mathbf{j} , also of unit length, pointing in the position of the positive y axis. Then, for any point P with coordinates (x, y) in the plane, the vector with initial point at the center O of the coordinate system and terminal point at P , one can write

$$\overrightarrow{OP} = x\mathbf{i} + y\mathbf{j} = \mathbf{i}x + \mathbf{j}y = (\mathbf{i}, \mathbf{j}) \begin{pmatrix} x \\ y \end{pmatrix}.$$

The vector \overrightarrow{OP} is called the position vector of the point P . Denote by E_2 the space of all position vectors in the plane.³ The family $\mathcal{X} = (\mathbf{i}, \mathbf{j})$ is clearly a basis of E_2 . The displayed equation above can also be expressed by saying that

$$\mathbf{R}_{\mathcal{X}} \overrightarrow{OP} = (x, y)^T,$$

where the superscript T denotes transpose; in other words, the representation of the vector \overrightarrow{OP} in the basis \mathcal{X} is the column vector $(x, y)^T$.

¹Notes for Course Mathematics 10.1 at Brooklyn College of CUNY. Attila Máté, November 7, 2007. Revised on April 15, 2013.

²Here ^{1, 2, ...} attached to vectors indicate superscripts, and are not exponents. As usual, in a field we assume that the multiplication is commutative. A field with a noncommutative multiplication is now called a *skew field* or a *division ring*.

³I.e., E_2 is the set of all vectors in the given plane with initial point at the center O of the given Cartesian coordinate system.

Consider the mapping T_ϕ such that, for each vector \overrightarrow{OP} , the result $T_\phi\overrightarrow{OP}$ is the vector \overrightarrow{OP} rotated about the point O by the angle ϕ (counterclockwise if $\phi \geq 0$, clockwise if $\phi < 0$). It is easy to see that T_ϕ is a linear transformation. Indeed, one can think of the transformation T_ϕ as rotating the whole plane by the angle ϕ about the fixed point O . If the vector $\vec{a} \in E_2$ gets rotated to \vec{a}' , then, for any real λ , $\lambda\vec{a}$ gets rotated to $\lambda\vec{a}'$; this shows that $T_\phi\lambda\vec{a} = \lambda T_\phi\vec{a}$. If the triangle OAB expresses the vector addition $\overrightarrow{OB} = \overrightarrow{OA} + \overrightarrow{AB}$,⁴ then the rotated triangle expresses the vector addition $T_\phi\overrightarrow{OB} = T_\phi\overrightarrow{OA} + T_\phi\overrightarrow{AB}$; this shows that for any two vectors $\vec{a}, \vec{b} \in E_2$ we have $T_\phi(\vec{a} + \vec{b}) = T_\phi\vec{a} + T_\phi\vec{b}$. Hence T_ϕ is indeed a linear transformation of E_2 into itself.

Simple geometric considerations show that

$$T_\phi\mathbf{i} = \mathbf{i} \cos \phi + \mathbf{j} \sin \phi = (\mathbf{i}, \mathbf{j}) \begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix} \quad \text{and} \quad T_\phi\mathbf{j} = -\mathbf{i} \sin \phi + \mathbf{j} \cos \phi = (\mathbf{i}, \mathbf{j}) \begin{pmatrix} -\sin \phi \\ \cos \phi \end{pmatrix}.$$

Using the linearity of T_ϕ , we therefore obtain that for any vector $\mathbf{v} = (x, y)^T \in \mathbb{R}_{2,1}$ we have

$$\begin{aligned} T_\phi\mathcal{X}\mathbf{v} &= T_\phi(\mathbf{i}, \mathbf{j}) \begin{pmatrix} x \\ y \end{pmatrix} = T_\phi(\mathbf{i}x + \mathbf{j}y) = xT_\phi\mathbf{i} + yT_\phi\mathbf{j} = (\mathbf{i}, \mathbf{j}) \begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix} x + (\mathbf{i}, \mathbf{j}) \begin{pmatrix} -\sin \phi \\ \cos \phi \end{pmatrix} y \\ &= (\mathbf{i}, \mathbf{j}) \left(\begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix} x + \begin{pmatrix} -\sin \phi \\ \cos \phi \end{pmatrix} y \right) = (\mathbf{i}, \mathbf{j}) \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \mathcal{X}A_\phi\mathbf{v}, \end{aligned}$$

where

$$A_\phi = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}.$$

That is, for any $\mathbf{v} \in \mathbb{R}_{2,1}$ we have $T_\phi\mathcal{X}\mathbf{v} = \mathcal{X}A_\phi\mathbf{v}$, or else, for the representation of the linear transformation T_ϕ we have $\mathbf{R}_{\mathcal{X}\mathcal{X}}T_\phi = A_\phi$.

Now, if we rotate the vector $\vec{a} \in E_2$ by an angle β to form $T_\beta\vec{a}$, and then we rotate it by the angle α to obtain the vector $T_\alpha T_\beta\vec{a}$, we can also arrive at the same vector by rotating the vector \vec{a} by the angle $\alpha + \beta$ to obtain $T_{\alpha+\beta}\vec{a}$. That is, for any vector $\vec{a} \in E_2$ we have $T_\alpha T_\beta\vec{a} = T_{\alpha+\beta}\vec{a}$; i.e., $T_\alpha T_\beta = T_{\alpha+\beta}$. Hence, for any vector $\mathbf{v} \in \mathbb{R}_{2,1}$, we have

$$T_\alpha T_\beta\mathcal{X}\mathbf{v} = T_\alpha(T_\beta\mathcal{X}\mathbf{v}) = T_\alpha(\mathcal{X}A_\beta\mathbf{v}) = T_\alpha\mathcal{X}(A_\beta\mathbf{v}) = \mathcal{X}A_\alpha(A_\beta\mathbf{v}) = \mathcal{X}(A_\alpha A_\beta)\mathbf{v}$$

on the one hand, and, on the other hand

$$T_\alpha T_\beta\mathcal{X}\mathbf{v} = T_{\alpha+\beta}\mathcal{X}\mathbf{v} = \mathcal{X}A_{\alpha+\beta}\mathbf{v}.$$

Thus, we have $\mathbf{R}_{\mathcal{X}\mathcal{X}}T_\alpha T_\beta = A_\alpha A_\beta$, and also $\mathbf{R}_{\mathcal{X}\mathcal{X}}T_{\alpha+\beta} = A_{\alpha+\beta}$. Since the matrix representation of a linear operator with respect to given bases is unique, we have $A_{\alpha+\beta} = A_\alpha A_\beta$. That is

$$\begin{aligned} \begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix} &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix} \\ &= \begin{pmatrix} \cos \alpha \cos \beta - \sin \alpha \sin \beta & -\cos \alpha \sin \beta - \sin \alpha \cos \beta \\ \sin \alpha \cos \beta + \cos \alpha \sin \beta & -\sin \alpha \sin \beta + \cos \alpha \cos \beta \end{pmatrix}; \end{aligned}$$

the second equality was obtained by multiplying out the matrices in the middle. Since the two matrices at the ends are equal, their corresponding entries must be equal. Hence, by the equalities of the entries in the first columns, we obtain

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta \quad \text{and} \quad \sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta.$$

Thus, the above considerations with linear transformations lead to a proof of the basic addition formulas for sine and cosine.

⁴Of course, \overrightarrow{AB} is not a position vector unless the point A is the same as O ; however, it is considered to be equal to a position vector of the same length and same direction.

Rotation of the coordinate system. If we rotate the coordinate vectors \mathbf{i} and \mathbf{j} to obtain $\mathbf{i}_\phi = T_\phi \mathbf{i}$ and $\mathbf{j}_\phi = T_\phi \mathbf{j}$, the family $\mathcal{Y} = (\mathbf{i}_\phi, \mathbf{j}_\phi)$ will also be a basis of the space E_2 of plane position vectors, and the above equations can also be written as $\mathcal{Y} = T_\phi \mathcal{X}$, where $\mathcal{X} = (\mathbf{i}, \mathbf{j})$, as above. Note $T_{-\phi} T_\phi \vec{a} = \vec{a} = T_\phi T_{-\phi} \vec{a}$ for any vector $\vec{a} \in E_2$. Indeed, the first of these equations just expresses the fact that if we rotate the vector \vec{a} first by an angle ϕ , and then rotate it by the angle $-\phi$, we get back the vector \vec{a} ; the second equation expresses a similar fact when we make the first rotation by the angle $-\phi$ and the second one by ϕ . Hence, we have $T_{-\phi} T_\phi = I = T_\phi T_{-\phi}$, where I is the identity transformation on E_2 . That is, we have $T_\phi^{-1} = T_{-\phi}$. I.e., for any column vector $\mathbf{v} \in \mathbb{R}_{2,1}$ we have

$$\mathcal{X}\mathbf{v} = I\mathcal{X}\mathbf{v} = (T_\phi T_{-\phi})\mathcal{X}\mathbf{v} = T_\phi(T_{-\phi}\mathcal{X}\mathbf{v}) = T_\phi\mathcal{X}A_{-\phi}\mathbf{v} = (T_\phi\mathcal{X})A_{-\phi}\mathbf{v} = \mathcal{Y}A_{-\phi}\mathbf{v}.$$

The equation $I\mathcal{X}\mathbf{v} = \mathcal{Y}A_{-\phi}\mathbf{v}$ can also be written as $\mathbf{R}_{\mathcal{Y}\mathcal{X}}I = A_{-\phi}$.

Observe that $A_{-\phi}$ is the inverse of A_ϕ . This is because

$$\mathcal{X}\mathbf{v} = T_\phi(T_{-\phi}\mathcal{X}\mathbf{v}) = T_\phi(\mathcal{X}A_{-\phi}\mathbf{v}) = T_\phi\mathcal{X}(A_{-\phi}\mathbf{v}) = \mathcal{X}A_\phi(A_{-\phi}\mathbf{v}) = \mathcal{X}A_\phi A_{-\phi}\mathbf{v},$$

Similarly, $\mathcal{X}\mathbf{v} = \mathcal{X}A_{-\phi}A_\phi\mathbf{v}$. Since these equations hold for any $\mathbf{v} \in \mathbb{R}_{2,1}$, they mean that the identity transformation is represented by the matrices $\mathbf{R}_{\mathcal{X}\mathcal{X}}I = A_\phi A_{-\phi} = A_{-\phi} A_\phi$. The equation $I\mathcal{X}\mathbf{v} = \mathcal{X}I_2\mathbf{v}$, where I_2 denotes the 2×2 identity matrix, shows that we also have $\mathbf{R}_{\mathcal{X}\mathcal{X}}I = I_2$. Thus $A_\phi A_{-\phi} = A_{-\phi} A_\phi = I_2$, so, indeed, $A_\phi^{-1} = A_{-\phi}$.⁵

Now, let $\mathbf{u} \in \mathbb{R}_{2,1}$ be an arbitrary vector, and let $\mathbf{v} = A_\phi\mathbf{u}$. Then $\mathbf{u} = A_{-\phi}\mathbf{v}$, since $A_{-\phi}$ is the inverse of the matrix A_ϕ . Thus, the equation $\mathcal{X}\mathbf{v} = \mathcal{Y}A_{-\phi}\mathbf{v}$ can be rewritten as $\mathcal{Y}\mathbf{u} = \mathcal{X}A_\phi\mathbf{u}$, or as $I\mathcal{Y}\mathbf{u} = \mathcal{X}A_\phi\mathbf{u}$. This equation can also be written as $\mathbf{R}_{\mathcal{X}\mathcal{Y}}I = A_\phi$.

Let P be a point in the plane, and let $\mathbf{R}_\mathcal{X}\overrightarrow{OP} = (x, y)^T$, and $\mathbf{R}_\mathcal{Y}\overrightarrow{OP} = (x_\phi, y_\phi)^T$. This means that, in the coordinate system C with the positive x axis pointing in the direction of the vector \mathbf{i} and the positive y axis pointing in the direction of the vector \mathbf{j} , the coordinates of the point P are (x, y) , and in the coordinate system C_ϕ with the same center but the axes rotated by an angle ϕ , the coordinates of P are (x_ϕ, y_ϕ) . We have

$$\overrightarrow{OP} = \mathcal{X} \begin{pmatrix} x \\ y \end{pmatrix} = \mathcal{Y} \begin{pmatrix} x_\phi \\ y_\phi \end{pmatrix}.$$

The equation $\mathcal{X}\mathbf{v} = \mathcal{Y}A_{-\phi}\mathbf{v}$ with $\mathbf{v} = (x, y)^T$ then implies that

$$\begin{aligned} \begin{pmatrix} x_\phi \\ y_\phi \end{pmatrix} &= A_{-\phi} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(-\phi) & -\sin(-\phi) \\ \sin(-\phi) & \cos(-\phi) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \cos \phi + y \sin \phi \\ -x \sin \phi + y \cos \phi \end{pmatrix}. \end{aligned}$$

This means that $x_\phi = x \cos \phi + y \sin \phi$ and $y_\phi = -x \sin \phi + y \cos \phi$, describing the way how the coordinates of a point transform under the rotation of a coordinate system by an angle ϕ . Similarly, the equation $\mathcal{Y}\mathbf{u} = \mathcal{X}A_\phi\mathbf{u}$ implies with $\mathbf{u} = (x_\phi, y_\phi)^T$

$$\begin{pmatrix} x \\ y \end{pmatrix} = A_\phi \begin{pmatrix} x_\phi \\ y_\phi \end{pmatrix} = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} x_\phi \\ y_\phi \end{pmatrix} = \begin{pmatrix} x_\phi \cos \phi - y_\phi \sin \phi \\ x_\phi \sin \phi + y_\phi \cos \phi \end{pmatrix}.$$

⁵The equation $A_\phi A_{-\phi} = I_2$ can be written as

$$\begin{aligned} A_\phi A_{-\phi} &= \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} \cos(-\phi) & -\sin(-\phi) \\ \sin(-\phi) & \cos(-\phi) \end{pmatrix} = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix} \\ &= \begin{pmatrix} \cos^2 \phi + \sin^2 \phi & 0 \\ 0 & \sin^2 \phi + \cos^2 \phi \end{pmatrix} = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

The equality of the last two matrices implies that $\cos^2 \phi + \sin^2 \phi = 1$. This is of course a well known identity; however, the point is that the argument described here gives a proof of this identity. A proof of this identity was already implicitly contained in an argument given earlier, since this identity also follows by applying the addition formula for cosine to the left-hand side of the equation $\cos(\phi + (-\phi)) = \cos 0$ and noting that $\cos 0 = 1$.

Equating the corresponding entries, this gives the coordinate transformation equations $x = x_\phi \cos \phi - y_\phi \sin \phi$ and $y = x_\phi \sin \phi + y_\phi \cos \phi$.

As an application, take $\phi = \pi/4$, denote by (x, y) the coordinates of a point by P in the original coordinate system C , and by (ξ, η) the coordinates of the same point in the coordinate system C' with axes rotated by $\pi/4$ (i.e., by 45° counterclockwise). Then, noting that $\cos(\pi/4) = \sin(\pi/4) = 1/\sqrt{2}$, the coordinate transformation equations with ξ and η replacing x_ϕ and y_ϕ become

$$\begin{aligned} \xi &= \frac{x+y}{\sqrt{2}}, & \text{and} & & x &= \frac{\xi-\eta}{\sqrt{2}}, \\ \eta &= \frac{-x+y}{\sqrt{2}}, & & & y &= \frac{\xi+\eta}{\sqrt{2}}. \end{aligned}$$

The equation

$$\frac{\xi^2}{2} - \frac{\eta^2}{2} = 1$$

represents a hyperbola with its real axis along the ξ coordinate axis (i.e., the first coordinate axis in the coordinate system C'), and imaginary axis along the η axis (i.e., the second axis in the coordinate system C'); its foci, described in the coordinate system C' , that is, in the ξ, η -coordinate system, are $(2, 0)$ and $(-2, 0)$.⁶ Using the coordinate transformation equations in the first column, the equation of the hyperbola can be written as

$$\frac{\left(\frac{x+y}{\sqrt{2}}\right)^2}{2} - \frac{\left(\frac{-x+y}{\sqrt{2}}\right)^2}{2} = 1,$$

or else as

$$xy = 1.$$

This is the equation of the hyperbola in question in the x, y -coordinate system. Using the coordinate transformation equations in the second column, the (x, y) coordinates of the focus with (ξ, η) coordinates $\xi = 2$ and $\eta = 0$ are $x = (2 - 0)/\sqrt{2} = 2/\sqrt{2} = \sqrt{2}$ and $y = (2 + 0)/\sqrt{2} = 2/\sqrt{2} = \sqrt{2}$. Similarly, the (x, y) coordinates of the focus with (ξ, η) coordinates $\xi = -2$ and $\eta = 0$ are $x = (-2 - 0)/\sqrt{2} = -2/\sqrt{2} = -\sqrt{2}$ and $y = (-2 + 0)/\sqrt{2} = -2/\sqrt{2} = -\sqrt{2}$. That is, the foci of the above hyperbola are located at the points $(\sqrt{2}, \sqrt{2})$ and $(-\sqrt{2}, -\sqrt{2})$, as described in the x, y -coordinate system.

REFERENCE

1. Hans Schneider–George Philip Barker, *Matrices and Linear Algebra*, Dover Publications, Inc., New York, 1989.

⁶The canonical equation of the hyperbola with center at the origin, with semi-real axis a in the direction of the coordinate axis ξ , and with semi-imaginary axis b , is

$$\frac{\xi^2}{a^2} - \frac{\eta^2}{b^2} = 1;$$

the foci are located at the points $(c, 0)$ and $(-c, 0)$ in the ξ, η -coordinate system, where $c^2 = a^2 + b^2$ ($c > 0$). In the present case, $a = b = \sqrt{2}$, so $c = 2$.