

## 1. Visualizing Matrices as Operations

This problem is going to help you visualize matrices as operations. For example, when we multiply a vector by a “rotation matrix,” we will see it “rotate” in the true sense here. Similarly, when we multiply a vector by a “reflection matrix,” we will see it be “reflected.” The way we will see this is by applying the operation to all the vertices of a polygon and seeing how the polygon changes.

Your TA will now show you how a unit square can be rotated, scaled, or reflected using matrices!

### Part 1: Rotation Matrices as Rotations

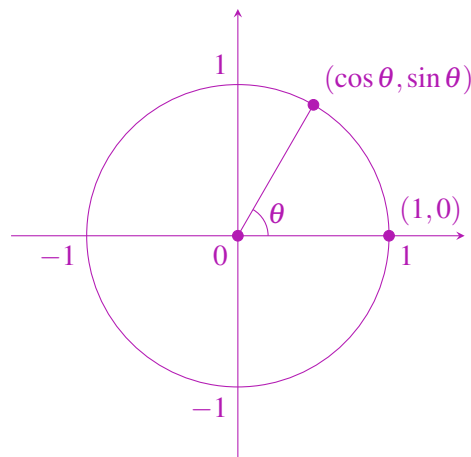
- We are given matrices  $\mathbf{T}_1$  and  $\mathbf{T}_2$ , and we are told that they will rotate the unit square by  $15^\circ$  and  $30^\circ$ , respectively. Design a procedure to rotate the unit square by  $45^\circ$  using only  $\mathbf{T}_1$  and  $\mathbf{T}_2$ , and plot the result in the IPython notebook. How would you rotate the square by  $60^\circ$ ?
- Try to rotate the unit square by  $60^\circ$  using only one matrix. What does this matrix look like?
- $\mathbf{T}_1$ ,  $\mathbf{T}_2$ , and the matrix you used in part (b) are called “rotation matrices.” They rotate any vector by an angle  $\theta$ . Show that a rotation matrix has the following form:

$$\mathbf{R} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

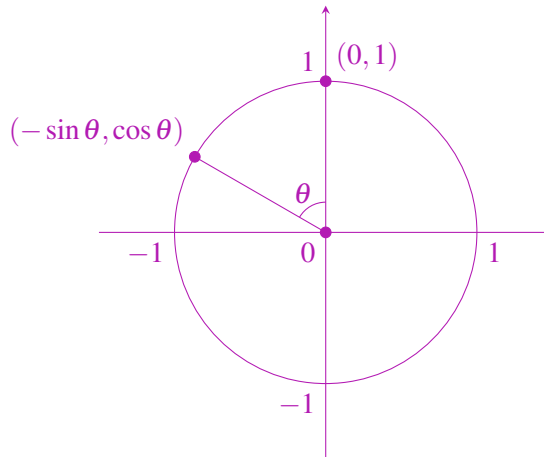
where  $\theta$  is the angle of rotation. (*Hint: Use your trigonometric identities!*)

**Answer:**

Let’s try to derive this matrix using trigonometry. Suppose we want to rotate the vector  $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$  by  $\theta$ .



We can use basic trigonometric relationships to see that  $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$  rotated by  $\theta$  becomes  $\begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}$ . Similarly, rotating the vector  $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$  by  $\theta$  becomes  $\begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix}$ :



We can also scale these pre-rotated vectors to any length we want,  $\begin{bmatrix} x \\ 0 \end{bmatrix}$  and  $\begin{bmatrix} 0 \\ y \end{bmatrix}$ , and we can observe graphically that they rotate to  $\begin{bmatrix} x \cos \theta \\ x \sin \theta \end{bmatrix}$  and  $\begin{bmatrix} -y \sin \theta \\ y \cos \theta \end{bmatrix}$ , respectively. Rotating a vector solely in the  $x$ -direction produces a vector with both  $x$  and  $y$  components, and, likewise, rotating a vector solely in the  $y$ -direction produces a vector with both  $x$  and  $y$  components.

Finally, if we want to rotate an arbitrary vector  $\begin{bmatrix} x \\ y \end{bmatrix}$ , we can combine what we derived above. Let  $x'$  and  $y'$  be the  $x$  and  $y$  components after rotation.  $x'$  has contributions from both  $x$  and  $y$ :  $x' = x \cos \theta - y \sin \theta$ . Similarly,  $y'$  has contributions from both components as well:  $y' = x \sin \theta + y \cos \theta$ . Expressing this in matrix form:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \cos \theta - y \sin \theta \\ x \sin \theta + y \cos \theta \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Thus, we've derived the 2-dimensional rotation matrix.

**Alternative solution:**

The reason the matrix is called a rotation matrix is because it translates the unit vector  $\begin{bmatrix} \cos \alpha \\ \sin \alpha \end{bmatrix}$  to give  $\begin{bmatrix} \cos(\alpha + \theta) \\ \sin(\alpha + \theta) \end{bmatrix}$ .

Proof:

$$\begin{aligned} \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \cos \alpha \\ \sin \alpha \end{bmatrix} &= \cos \alpha \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix} + \sin \alpha \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix} \\ &= \begin{bmatrix} \cos \alpha \cos \theta - \sin \alpha \sin \theta \\ \cos \alpha \sin \theta + \sin \alpha \cos \theta \end{bmatrix} \\ &= \begin{bmatrix} \cos(\alpha + \theta) \\ \sin(\alpha + \theta) \end{bmatrix} \end{aligned}$$

- (d) Now, we want to get back the original unit square from the rotated square in part (b). What matrix should we use to do this? *Don't use inverses!*

**Answer:**

Use a rotation matrix that rotates by  $-60^\circ$ .

$$\begin{bmatrix} \cos(-60^\circ) & -\sin(-60^\circ) \\ \sin(-60^\circ) & \cos(-60^\circ) \end{bmatrix}$$

- (e) Use part (d) to obtain the “inverse” rotation matrix for a matrix that rotates a vector by  $\theta$ . Multiply the inverse rotation matrix with the rotation matrix and vice-versa. What do you get?

**Answer:**

The inverse matrix is as follows:

$$\begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$$

We can see from this inverse matrix that the product of the rotation matrix and its inverse is the identity matrix.

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

## Part 2: Commutativity of Operations

A natural question to ask is the following: Does the *order* in which you apply these operations matter? Follow your TA to obtain the answers to the following questions!

- Let's see what happens to the unit square when we rotate the matrix by  $60^\circ$  and then reflect it along the  $y$ -axis.
- Now, let's see what happens to the unit square when we first reflect it along the  $y$ -axis and then rotate the matrix by  $60^\circ$ .
- Try to do steps (a) and (b) by multiplying the reflection and rotation matrices together (in the correct order for each case). What does this tell you?
- If you reflected the unit square twice (along any pair of axes), do you think the order in which you applied the reflections would matter? Why/why not?

**Answer:**

It turns out that reflections are not commutative unless the two reflection axes are perpendicular to each other. For example, if you reflect about the  $x$ -axis and the  $y$ -axis, it is commutative. But if you reflect about the  $x$ -axis and  $x = y$ , it is not commutative.

## 2. Proofs

- (a) Suppose for some non-zero vector  $\vec{x}$ ,  $\mathbf{A}\vec{x} = \vec{0}$ . Prove that the columns of  $\mathbf{A}$  are linearly dependent.

**Answer:**

Begin by defining column vectors  $\vec{v}_1 \dots \vec{v}_n$ .

$$\mathbf{A} = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}$$

Thus, we can represent the multiplication  $\mathbf{A}\vec{x}$  as

$$\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} | \\ \vec{x} \\ | \end{bmatrix} = \sum x_i \vec{v}_i = \vec{0}$$

Note that the equation above is the definition of linear dependence. That is, there exist coefficients, at least one which is non-zero, such that the sum of the vectors weighted by the coefficients is zero. These coefficients are the elements of the non-zero vector  $\vec{x}$ .

- (b) Suppose there exists a matrix  $\mathbf{A}$  whose columns are linearly dependent. Prove that if there exists a solution to  $\mathbf{A}\vec{x} = \vec{b}$ , then there are infinitely many solutions. What is the physical interpretation of this statement?

**Answer:**

Begin by defining column vectors  $\vec{v}_1, \dots, \vec{v}_n$ .

$$\mathbf{A} = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}$$

Recall the definition of linear dependence:

$$\sum \alpha_i \vec{v}_i = \vec{0} \quad \exists i, \alpha_i \neq 0$$

Note the constraint that not all of the weights  $\alpha_i$  can equal 0! (Equivalently, at least one of the weights must be non-zero.) This is extremely important—overlooking this detail will make the proof incorrect. What does this imply? It implies that there exists some  $\vec{\alpha}$  such that  $\mathbf{A}\vec{\alpha} = \vec{0}$ , so that for any  $\vec{x}$ , where  $\mathbf{A}\vec{x} = \vec{b}$ , then  $(\vec{x} + k\vec{\alpha}), \forall k \in \mathbb{R}$ , is also a valid solution.

$$\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} | \\ \vec{x} \\ | \end{bmatrix} = \begin{bmatrix} | \\ \vec{b} \\ | \end{bmatrix}$$

$$\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} \left( \begin{bmatrix} | \\ \vec{x} \\ | \end{bmatrix} + \begin{bmatrix} | \\ \vec{\alpha} \\ | \end{bmatrix} \right) = \begin{bmatrix} | \\ \vec{b} \\ | \end{bmatrix} + \begin{bmatrix} | \\ 0 \\ | \end{bmatrix}$$

Therefore, if a solution  $\vec{x}$  exists, infinite solutions must exist:  $\exists \vec{x}, \mathbf{A}\vec{x} = \vec{b} \iff \mathbf{A}(\vec{x} + k\vec{\alpha}) = \vec{b}, \forall k \in \mathbb{R}$ .

- (c) Suppose we have an experiment where we have  $n$  measurements of linear combinations of  $n$  unknowns. We want to show that if at least one of the experiment's measurements can be predicted from the other measurements, then there will be either infinite or no solutions. Reword this statement into a proof problem and, as practice, complete the proof.

**Answer:**

**Reworded:** Prove that if an  $n \times n$  matrix  $\mathbf{A}$  has rows that are linearly dependent, there will be either infinite or no solutions to  $\mathbf{A}\vec{x} = \vec{b}$ .

Define row vectors  $\vec{v}_1, \dots, \vec{v}_n$ .

$$\mathbf{A} = \begin{bmatrix} - & \vec{v}_1 & - \\ - & \vec{v}_2 & - \\ - & \vdots & - \\ - & \vec{v}_n & - \end{bmatrix}$$

Recall the definition of linear dependence. Note that the summation is not over all  $i \in 1, \dots, n$  but only in the set where  $i \neq j$ ! This is extremely important. If this detail is overlooked, the proof will be incorrect.

$$\vec{v}_j = \sum_{i \neq j} \alpha_i \vec{v}_i$$

Therefore, we know we can row reduce this matrix to have a row of zeros (by definition of Gaussian elimination).

$$\begin{bmatrix} - & \vec{v}_1' & - \\ - & \vec{v}_2' & - \\ - & \vdots & - \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} | \\ | \\ | \\ | \end{bmatrix} \vec{x} = \begin{bmatrix} | \\ | \\ | \\ b_n \end{bmatrix}$$

We know that if  $b_n = 0$ , then  $\vec{x}$  has infinite solutions. If  $b_n \neq 0$ , then the system is inconsistent and has no solutions.

- (d) **Practice Problem:** Now suppose there exist two unique vectors  $\vec{x}_1$  and  $\vec{x}_2$  that both satisfy  $\mathbf{A}\vec{x} = \vec{b}$ , that is,  $\mathbf{A}\vec{x}_1 = \vec{b}$  and  $\mathbf{A}\vec{x}_2 = \vec{b}$ . Prove that the columns of  $\mathbf{A}$  are linearly dependent.

**Answer:**

Let us consider the difference of the two equations:

$$\mathbf{A}\vec{x}_1 - \mathbf{A}\vec{x}_2 = \mathbf{A}(\vec{x}_1 - \vec{x}_2) = \vec{b} - \vec{b} = \vec{0}$$

Once again, we've reached the definition of linear dependence since  $\vec{x}_1 - \vec{x}_2 \neq \vec{0}$ .

- (e) **Challenging Practice Problem:** Prove that for a  $m \times n$  matrix, the number of linearly independent vectors (both column and row) is at most  $\min(m, n)$ .

**Answer:**

The number of pivots can, at most, be equal to  $\min(m, n)$ , implying that  $\text{col rank} = \text{row rank}$ .