EECS 16A Designing Information Devices and Systems I
Fall 2019 Discussion 2B

There are two different views of matrix-vector multiplication. Let $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2}$. The column view of matrix $\mathbf{A}$ can be defined as $\begin{bmatrix} \vec{a}_1 & \vec{a}_2 \end{bmatrix}$, where $\vec{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} \in \mathbb{R}^2$ and $\vec{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} \in \mathbb{R}^2$. The row view of $\mathbf{A}$ is $\begin{bmatrix} \vec{a}_1' \\ \vec{a}_2' \end{bmatrix}$, where $\vec{a}_1' = \begin{bmatrix} a_{11} & a_{12} \end{bmatrix} \in \mathbb{R}^2$ and $\vec{a}_2' = \begin{bmatrix} a_{21} & a_{22} \end{bmatrix} \in \mathbb{R}^2$. If $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \in \mathbb{R}^2$ then the column view of matrix-vector multiplication computes $\mathbf{A}\vec{b}$ as a linear combination of the columns of $\mathbf{A}$,

$$\mathbf{A}\vec{b} = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 \end{bmatrix} \vec{b} = b_1\vec{a}_1 + b_2\vec{a}_2 = \begin{bmatrix} a_{11}b_1 + a_{12}b_2 \\ a_{21}b_1 + a_{22}b_2 \end{bmatrix}. \tag{1}$$

The row view of matrix-vector multiplication computes $\mathbf{A}\vec{b}$ as a series of inner products,

$$\mathbf{A}\vec{b} = \begin{bmatrix} \vec{a}_1' \\ \vec{a}_2' \end{bmatrix} \vec{b} = \begin{bmatrix} \vec{a}_1'\vec{b} \\ \vec{a}_2'\vec{b} \end{bmatrix} = \begin{bmatrix} a_{11}b_1 + a_{12}b_2 \\ a_{21}b_1 + a_{22}b_2 \end{bmatrix}. \tag{2}$$

1. Span Proofs

Given some set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$, show the following:

(a)

$$\text{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} = \text{span}\{\alpha\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}, \text{ where } \alpha \text{ is a non-zero scalar}$$

In other words, we can scale our spanning vectors and not change their span.

(b) **(Practice)**

$$\text{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} = \text{span}\{\vec{v}_2, \vec{v}_1, \ldots, \vec{v}_n\}$$

In other words, we can swap the order of our spanning vectors and not change their span.

Answer:

(a) Suppose we have some arbitrary $\vec{q} \in \text{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$. For some scalars a_i:

$$\vec{q} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n = \left(\frac{a_1}{\alpha}\right)\alpha\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n.$$

Scalar multiplication cancels out. Thus, we have shown that $\vec{q} \in \text{span}\{\alpha\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$. Therefore, we have $\text{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} \subseteq \text{span}\{\alpha\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$. Now, we must show the other direction. Suppose we have some arbitrary $\vec{r} \in \text{span}\{\alpha\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$. For some scalars b_i:

$$\vec{r} = b_1(\alpha\vec{v}_1) + b_2\vec{v}_2 + \cdots + b_n\vec{v}_n = (b_1\alpha)\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_n\vec{v}_n.$$

Thus, we have shown that $\vec{r} \in \text{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$. Therefore, we now have $\text{span}\{\alpha\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} \subseteq \text{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$. Combining this with the earlier result, the spans are thus the same.

(b) Suppose $\vec{q} \in \text{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$. For some scalars a_i:

$$\vec{q} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n = a_2\vec{v}_2 + a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$$

Swapping the order in addition does not affect the sum, so $\text{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} \subseteq \text{span}\{\vec{v}_2, \vec{v}_1, \ldots, \vec{v}_n\}$. Similarly, starting with some $\vec{r} \in \text{span}\{\vec{v}_2, \vec{v}_1, \ldots, \vec{v}_n\}$, again swapping the order does not affect the sum, so putting both together, the spans are thus the same.

2. 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

(a) We are given matrices $\mathbf{T}_1$ and $\mathbf{T}_2$, and we are told that they will rotate the unit square by 15° and 30°, respectively. Design a procedure to rotate the unit square by 45° 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°?

Answer:

Apply $\mathbf{T}_1$ and $\mathbf{T}_2$ in succession to rotate the unit square by 45°. To rotate the square by 60°, you can either apply $\mathbf{T}_2$ twice, or if you prefer variety, apply $\mathbf{T}_1$ twice and $\mathbf{T}_2$ once.

(b) Try to rotate the unit square by 60° using only one matrix. What does this matrix look like?

Answer: This matrix will look like the rotation matrix that rotates a vector by 60°. This matrix can be composed by multiplying $\mathbf{T}_1$ by $\mathbf{T}_1$ by $\mathbf{T}_2$ (or equivalently, $\mathbf{T}_2$ by $\mathbf{T}_2$).

(c) $\mathbf{T}_1$, $\mathbf{T}_2$, and the matrix you used in part (b) are called "rotation matrices." They rotate any vector by an angle θ. Show that a rotation matrix has the following form:

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

where θ is the angle of rotation. To do this consider rotating the unit vector $\begin{bmatrix} cos(\alpha) \\ sin(\alpha) \end{bmatrix}$ by θ degrees using the matrix $\mathbf{R}$.

(**Definition:** A vector, $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \end{bmatrix}$, is a unit vector if $\sqrt{v_1^2 + v_2^2 + \ldots} = 1$.)

(*Hint: Use your trigonometric identities!*)

Answer:

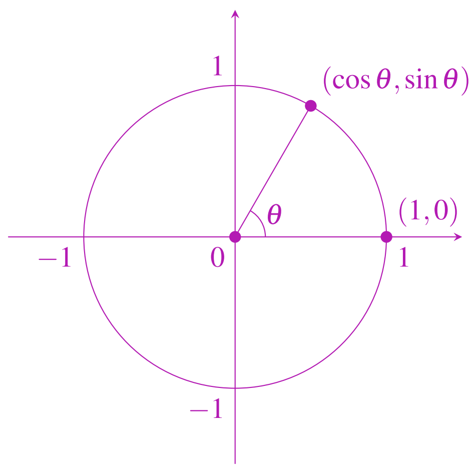
The reason the matrix is called a rotation matrix is because it transforms the unit vector $\begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}$ to give $\begin{bmatrix} \cos(\alpha+\theta) \\ \sin(\alpha+\theta) \end{bmatrix}$.

Proof:

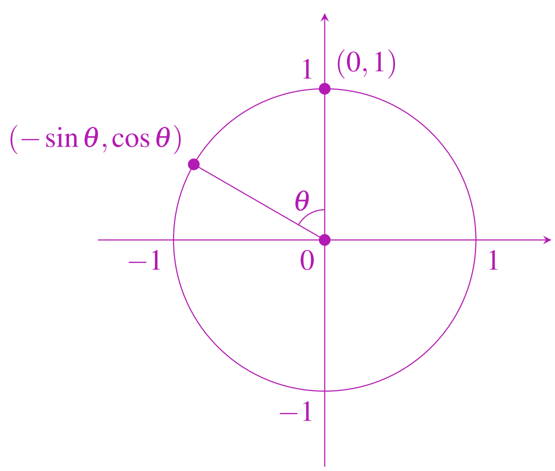
$$\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}\cos\alpha\\ \sin\alpha\end{bmatrix} = \cos\alpha\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix} + \sin\alpha\begin{bmatrix}-\sin\theta\\ \cos\theta\end{bmatrix}$$
$$= \begin{bmatrix}\cos\alpha\cos\theta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

Alternative solution:

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



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

(d) **(Practice)** 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? (**Note:** Don't use inverses! Answer this question using your intuition, we will visit inverses very soon in lecture!)

Answer:

Use a rotation matrix that rotates by $-60°$.

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

(e) **(Practice)** Use part (d) to obtain the "inverse" rotation matrix for a matrix that rotates a vector by θ. 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 that for any $\vec{v} \in \mathbb{R}^2$ that the product of the rotation matrix with $\vec{v}$ followed by the product of the inverse results in the original $\vec{v}$.

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \left(\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \vec{v} \right) = \vec{v}$$

(f) What are the matrices that reflect a vector about the (i) x-axis, (ii) y-axis, and (iii) $x = y$

Answer:

The matrix that reflects about the x-axis:

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

The matrix that reflects about the y-axis:

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

and the matrix that reflects about $x = y$:

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

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!

(a) Let's see what happens to the unit square when we rotate the square by 60° and then reflect it along the y-axis.

(b) Now, let's see what happens to the unit square when we first reflect the square along the y-axis and then rotate it by 60°.

Answer: (For parts (a) and (b)): The two operations are not the same.

(c) 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?

Answer:

The resulting matrices that are obtained (by multiplying the two matrices) are different depending on the order of multiplication.

(d) 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.

(Practice) Part 3: Distributivity of Operations

(a) The distributivity property of matrix-vector multiplication holds for any vectors and matrices. Show for general $\mathbf{A} \in \mathbb{R}^{2\times 2}$ and $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^2$ that $\mathbf{A}(\vec{v}_1 + \vec{v}_2) = \mathbf{A}\vec{v}_1 + \mathbf{A}\vec{v}_2$.

Answer: Matrix-vector multiplication distributes because scalar multiplication distributes.

$$\mathbf{A}(\vec{v}_1 + \vec{v}_2) = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 \end{bmatrix} (\vec{v}_1 + \vec{v}_2) \tag{3}$$

$$= (v_{11} + v_{21})\vec{a}_1 + (v_{12} + v_{22})\vec{a}_2 \tag{4}$$

$$= \begin{bmatrix} a_{11}(v_{11} + v_{21}) + a_{12}(v_{12} + v_{22}) \\ a_{21}(v_{11} + v_{21}) + a_{22}(v_{12} + v_{22}) \end{bmatrix} \tag{5}$$

$$= \begin{bmatrix} a_{11}v_{11} + a_{12}v_{12} \\ a_{21}v_{11} + a_{22}v_{12} \end{bmatrix} + \begin{bmatrix} a_{11}v_{21} + a_{12}v_{22} \\ a_{21}v_{21} + a_{22}v_{22} \end{bmatrix} \tag{6}$$

$$= \mathbf{A}\vec{v}_1 + \mathbf{A}\vec{v}_2 \tag{7}$$

