EECS 16A Designing Information Devices and Systems I
Fall 2021 Discussion 3B

1. Inverses

In general, the *inverse* of a matrix "undoes" the operation that a matrix performs. Mathematically, we write this as

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I},$$

where $\mathbf{A}^{-1}$ is the inverse of $\mathbf{A}$. Intuitively, this means that applying a matrix to a vector and then subsequently applying its inverse is the same as leaving the vector untouched.

Properties of Inverses

For a matrix $\mathbf{A}$, if its inverse exists, then:

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

$$\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$$

$$(k\mathbf{A})^{-1} = \tfrac{1}{k}\mathbf{A}^{-1} \qquad \text{for a nonzero scalar } k \in \mathbb{R}$$

$$\left(\mathbf{A}^{T}\right)^{-1} = \left(\mathbf{A}^{-1}\right)^{T} \qquad T \text{ is "Transpose"}$$

$$(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \qquad \text{assuming } \mathbf{A},\mathbf{B} \text{ are both invertible}$$

(a) Suppose $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are all invertible matrices.
Prove that $(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$.

Answer:

Matrix multiplication is *not commutative*, so you cannot flip the matrices around within the product. However, it is associative so that you can place parentheses freely:

$$\begin{aligned}
\mathbf{C}^{-1}\,\mathbf{B}^{-1}\,\mathbf{A}^{-1}\;\mathbf{A}\;\mathbf{B}\;\mathbf{C} &= \mathbf{C}^{-1}\,\mathbf{B}^{-1}\left(\mathbf{A}^{-1}\,\mathbf{A}\right)\mathbf{B}\,\mathbf{C} \\
&= \mathbf{C}^{-1}\,\mathbf{B}^{-1}\,\mathbf{I}\,\mathbf{B}\,\mathbf{C} \\
&= \mathbf{C}^{-1}\left(\mathbf{B}^{-1}\,\mathbf{B}\right)\mathbf{C} \\
&= \mathbf{C}^{-1}\,\mathbf{I}\,\mathbf{C} \\
&= \mathbf{C}^{-1}\,\mathbf{C} \\
&= \mathbf{I} \qquad \square
\end{aligned}$$

(b) Now consider the following four matrices.

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \qquad \mathbf{D} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

i. What do each of these matrices do when you multiply them by a vector $\vec{x}$? Draw a diagram.
ii. Intuitively, can these operations be undone? Why or why not? Make an intuitive argument.
iii. Are the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ invertible?
iv. Can you find anything in common about the rows (and columns) of $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$? (*Bonus:* How does this relate to the invertibility of $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$?)
v. Are all square matrices invertible?
vi. (PRACTICE) How can you find the inverse of a general $n \times n$ matrix?

Answer:

i.
- $\mathbf{A}$: Preserves the x component and sets the y component to zero.
- $\mathbf{B}$: Preserves the y component and sets the x component to zero.
- $\mathbf{C}$: Replaces the x and y components with the average of the x and y components.
- $\mathbf{D}$: Yields a weighted sum of x and y components. Places the sum in x and twice the sum in y.

ii. Intuitively, none of these operations can be undone because we lost some information. In the first two, we lost one component of the original. In the third case, we replaced both x and y with the average of the two. Thus, different inputs could lead to the same average and we wouldn't be to tell them apart. In the fourth case, we took a weighted sum of the x and y components. There are different values for x and y that could lead to the same sum. However, we cannot recover the original x and y because we didn't compute two unique weighted sums. Instead, we just multiplied the sum by two for the y component of the output.

iii. Since the operations are not one-to-one reversible, $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ are not invertible.

iv. The rows of $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ are all linearly dependent. The same is true for the columns. The generalization is that if a matrix is not invertible, then its rows and columns will be linearly dependent.

v. No. We have seen in the above parts that there are square matrices that are not invertible.

vi. We know that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. If we treat this as our now familiar $\mathbf{A}\vec{x} = \vec{b}$, we can use Gaussian elimination:

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{I} \end{array}\right] \Longrightarrow \left[\begin{array}{c|c} \mathbf{I} & \mathbf{A}^{-1} \end{array}\right]$$

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. Suggest some methods to rotate the unit square by 45° using only $\mathbf{T}_1$ and $\mathbf{T}_2$. How would you rotate the square by 60°? Your TA will show you the result in the iPython notebook.

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) Find a single matrix $\mathbf{T}_3$ to rotate the unit square by 60°. Your TA will show you the result in the iPython notebook.

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:

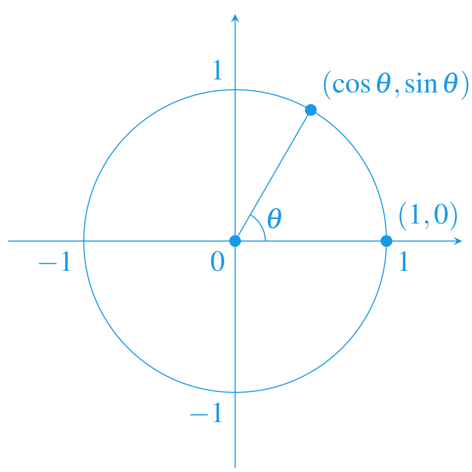
The reason the matrix is called a rotation matrix is because it transforms the unit vector $\begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}$ to give $\begin{bmatrix} \cos(\alpha+\theta) \\ \sin(\alpha+\theta) \end{bmatrix}$.

Proof:

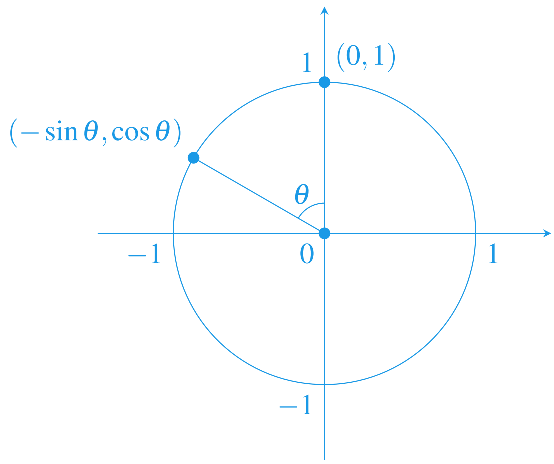
$$\begin{aligned}
\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix} &= \cos\alpha \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} + \sin\alpha \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} \\
&= \begin{bmatrix}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

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) 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) 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? Your TA will demonstrate parts (a) and (b) in the iPython notebook.

(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°. Is this the same as in part (a)?

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.

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.

$$
\begin{aligned}
\mathbf{A}(\vec{v}_1+\vec{v}_2) &= \begin{bmatrix}\vec{a}_1 & \vec{a}_2\end{bmatrix}(\vec{v}_1+\vec{v}_2) && (1)\\
&= (v_{11}+v_{21})\vec{a}_1+(v_{12}+v_{22})\vec{a}_2 && (2)\\
&= \begin{bmatrix}a_{11}(v_{11}+v_{21})+a_{12}(v_{12}+v_{22})\\ a_{21}(v_{11}+v_{21})+a_{22}(v_{12}+v_{22})\end{bmatrix} && (3)\\
&= \begin{bmatrix}a_{11}v_{11}+a_{12}v_{12}\\ a_{21}v_{11}+a_{22}v_{12}\end{bmatrix}+\begin{bmatrix}a_{11}v_{21}+a_{12}v_{22}\\ a_{21}v_{21}+a_{22}v_{22}\end{bmatrix} && (4)\\
&= \mathbf{A}\vec{v}_1+\mathbf{A}\vec{v}_2 && (5)
\end{aligned}
$$