EECS 16A Designing Information Devices and Systems I
Fall 2018 Discussion 2B

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

(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 $\mathbf{T}_2$).

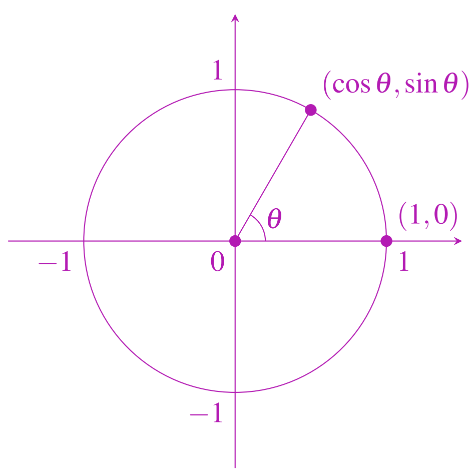
(c) $\mathbf{T}_1$, $\mathbf{T}_2$, and the matrix you used in part (b) are called "rotation matrices." They rotate any vector by an angle θ. Show that a rotation matrix has the following form:

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

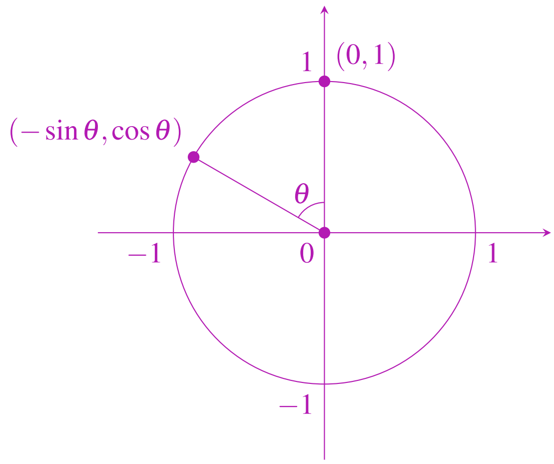
where θ is the angle of rotation. (*Hint: Use your trigonometric identities!*)

Answer:

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

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!* (**Note:** We do not expect you to know inverses at this point; we will cover them soon.)

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

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

[Practice Problem] 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.

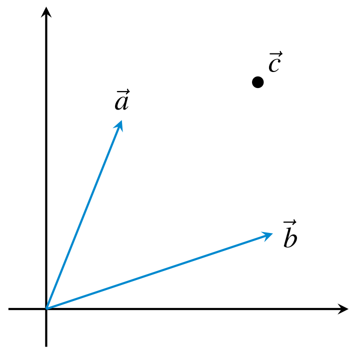
(d) If you reflected the unit square twice (along any pair of axes), do you think the order in which you applied the reflections would matter? Why/why not?

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.

2. Visualizing Span

We are given a point $\vec{c}$ that we want to get to, but we can only move in two directions: $\vec{a}$ and $\vec{b}$. We know that to get to $\vec{c}$, we can travel along $\vec{a}$ for some amount α, then change direction, and travel along $\vec{b}$ for some amount β. We want to find these two scalars α and β, such that we reach point $\vec{c}$. That is, $\alpha\vec{a} + \beta\vec{b} = \vec{c}$.



(a) First, consider the case where $\vec{a} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and $\vec{c} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$. Draw these vectors on a sheet of paper. Now find the two scalars α and β, such that we reach point $\vec{c}$. What are these scalars if we use $\vec{a} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ instead?

Answer: First set: $\alpha = -4, \beta = 2$
Second set: $\alpha = 6, \beta = -4$

(b) Now formulate the general problem as a system of linear equations and write it in matrix form.

Answer:

$$\begin{cases} \alpha a_1 + \beta b_1 = c_1 \\ \alpha a_2 + \beta b_2 = c_2 \end{cases}$$

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

3. Proofs

(a) **[Practice Problem]:** 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 $\vec{a}_1 \ldots \vec{a}_n$.

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

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

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

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) **[Practice Problem]:** 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 $\vec{x}_1 - \vec{x}_2 \neq \vec{0}$. We can apply the results from Part (a), setting $\vec{x} = \vec{x}_1 - \vec{x}_2$.

(c) 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.

Answer:

Begin by defining column vectors $\vec{a}_1, \ldots, \vec{a}_n$.

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

Recall the definition of linear dependence:

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

Note the constraint that not all of the weights α_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.

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

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

Therefore, if a solution $\vec{x}$ exists, infinite solutions must exist: $\exists \vec{x}, \mathbf{A}\vec{x} = \vec{b} \iff \mathbf{A}(\vec{x} + k\vec{\alpha}) = \vec{b}, \forall k \in \mathbb{R}$. Physically, taking the example of $\vec{x} \in \mathbb{R}^3$, the set of solutions is not just a single point but a line or plane.

4. Matrix Multiplication

Consider the following matrices:

$$\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 9 & 5 & 7 \\ 4 & 3 & 2 & 2 \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} 5 & 5 & 8 \\ 6 & 1 & 2 \\ 4 & 1 & 7 \\ 3 & 2 & 2 \end{bmatrix} \quad \mathbf{E} = \begin{bmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{bmatrix} \quad \mathbf{F} = \begin{bmatrix} 5 & 3 & 4 \\ 1 & 8 & 2 \\ 2 & 3 & 5 \end{bmatrix}$$

For each matrix multiplication problem, if the product exists, find the product by hand. Otherwise, explain why the product does not exist.

(a) **AB**

Answer: Since both $\mathbf{A}$ and $\mathbf{B}$ are 2×2 matrices, the product exists and is a 2×2 matrix.
$\mathbf{AB} = \begin{bmatrix} 11 & 6 \\ 12 & 7 \end{bmatrix}$.

(b) **BA**

Answer: Since both $\mathbf{A}$ and $\mathbf{B}$ are 2×2 matrices, the product exists and is a 2×2 matrix.
$\mathbf{BA} = \begin{bmatrix} 7 & 18 \\ 4 & 11 \end{bmatrix}$.

(c) **CD**

Answer: Since $\mathbf{C}$ is a 2×4 matrix and $\mathbf{D}$ is a 4×3 matrix, the product exists and is a 2×3 matrix.
$\mathbf{CD} = \begin{bmatrix} 100 & 33 & 75 \\ 52 & 29 & 56 \end{bmatrix}$.

(d) **DC**

Answer: Since $\mathbf{C}$ is a 2×4 matrix and $\mathbf{D}$ is a 4×3 matrix, the product does not exist. This is because the number of columns in the first matrix ($\mathbf{D}$) should match the number of rows in the second matrix ($\mathbf{C}$) for this product to be defined.

(e) **EF**

Answer: Since $\mathbf{E}$ and $\mathbf{F}$ are both 3×3 matrices, the product exists and is another 3×3 matrix.
$\mathbf{EF} = \begin{bmatrix} 53 & 50 & 64 \\ 34 & 70 & 57 \\ 33 & 90 & 44 \end{bmatrix}$.

(f) **FE**

Answer: Since $\mathbf{E}$ and $\mathbf{F}$ are both 3×3 matrices, the product exists and is another 3×3 matrix.
$\mathbf{FE} = \begin{bmatrix} 65 & 56 & 59 \\ 40 & 59 & 66 \\ 45 & 62 & 43 \end{bmatrix}$.