Lecture 14 - Linear Transformations and Matrix Multiplication

October 28, 2022

Goals: Characterize a linear function in terms of its interaction with vector addition and scalar multiplication, compute rotation matrices, and compose linear functions using matrix multiplication.

Recall the definition of linear and affine functions (in matrix form):

<u>**Definition:**</u> A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called

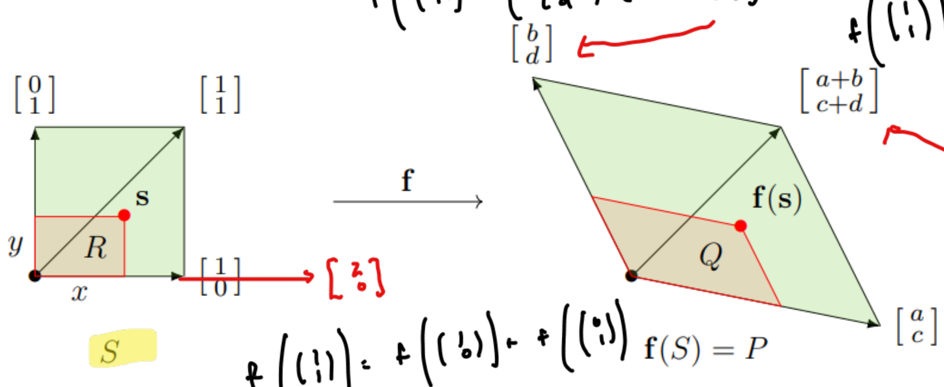
- a **linear function** or **linear transformation** if there is an $m \times n$ matrix A such that $f(\mathbf{x}) = A\mathbf{x}$ (recall that the jth column of A is $f(\mathbf{e}_j)$).
- an **affine function** or **affine transformation** if there is an $m \times n$ matrix A such that $f(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$.

Let us visualize the effect of the linear transformation $f : \mathbb{R}^2 \to \mathbb{R}^2$ with associated matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, where the columns are not multiples of each other. Let S denote the unit square $\{(x, y) : 0 \le x, y \le 1\}$. We can think of $\mathbb{R}^2$ as being tiled by copies of S. Let $f(S)$ denote the output of f on S (that is, the collection of points $f(\mathbf{s})$ with $\mathbf{s} \in S$). We call this the **image** of S under f.

$f\left(\begin{pmatrix}0\\1\end{pmatrix}\right) = \begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}b\\d\end{pmatrix}$

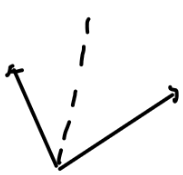
$f\left(\begin{pmatrix}1\\1\end{pmatrix}\right) = \begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}a+b\\c+d\end{pmatrix}$

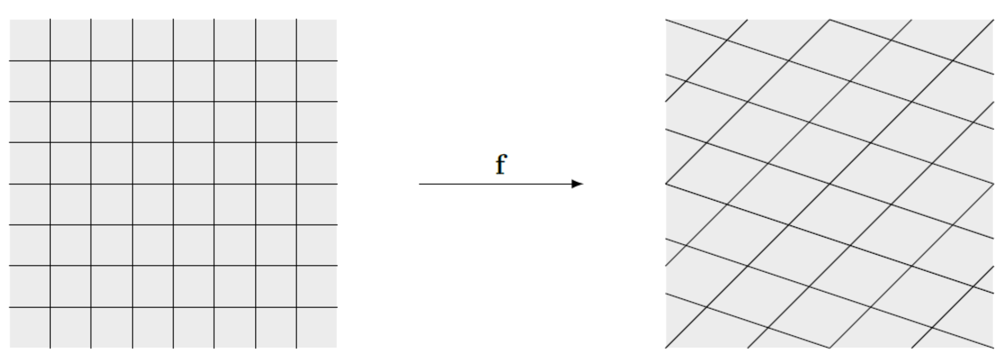
$f\left(\begin{bmatrix}2\\0\end{bmatrix}\right) = \begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}2\\0\end{pmatrix} = \begin{pmatrix}2a\\2c\end{pmatrix} = 2\begin{pmatrix}a\\c\end{pmatrix} = 2f\left(\begin{pmatrix}1\\0\end{pmatrix}\right)$



$f\left(\begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}a\\c\end{pmatrix}$

$f\left(\begin{pmatrix}1\\1\end{pmatrix}\right) = f\left(\begin{pmatrix}1\\0\end{pmatrix}\right) + f\left(\begin{pmatrix}0\\1\end{pmatrix}\right)$





works for all, not just standard basis

- **Linearity Principle:** for $c_1, c_2 \in \mathbb{R}$ and $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2$, we have $f(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1 f(\mathbf{v}_1) + c_2 f(\mathbf{v}_2)$. works w/ vec addition & scalar mult
- **Tiling Principle:** f transforms the tiling of $\mathbb{R}^2$ by copies of S into the tiling of $\mathbb{R}^2$ by copies of $f(S)$.

It turns that we can characterize linear functions based on their interaction with the vector operations.

Theorem 14.2.1: A function $g : \mathbb{R}^n \to \mathbb{R}^m$ is linear precisely when it respects the vector operations

$$f(c_1\vec{v}_1 + c_2\vec{v}_2) = \underset{\text{matrix}}{A}(c_1 v_1 + c_2 v_2)$$

- $g(c\mathbf{x}) = cg(\mathbf{x})$ scalar multiplication
- $g(\mathbf{x}+\mathbf{y}) = g(\mathbf{x}) + g(\mathbf{y})$ vector addition

for all scalars $c \in \mathbb{R}$ and vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

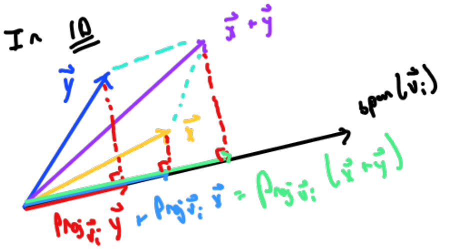
If $g : \mathbb{R}^n \to \mathbb{R}^m$ and $h : \mathbb{R}^p \to \mathbb{R}^n$ are linear, then so is the composition $g \circ h : \mathbb{R}^p \to \mathbb{R}^m$.

Example 1: For any non-zero linear subspace V of $\mathbb{R}^n$, the projection function $\mathbf{Proj}_V : \mathbb{R}^n \to \mathbb{R}^n$ is linear (We actually discussed this in the case where V is one dimensional in Lecture 6).

One dimensional: $\text{Proj}_{\vec{v}}\{c_1\vec{x}_1 + \dots + c_k\vec{x}_k\} = c_1 \text{Proj}_{\vec{v}}\vec{x}_1 + c_2 \text{Proj}_{\vec{v}}\vec{x}_2 + \dots + c_k \text{Proj}_{\vec{v}}\vec{x}_k$

Let $\{\vec{v}_1, \dots, \vec{v}_k\}$ be an orthogonal basis of V. By orthogonal projection theorem, $\text{Proj}_V \vec{v} = \text{Proj}_{\vec{v}_1}\vec{v} + \dots + \text{Proj}_{\vec{v}_k}\vec{v}$

In 1D



Algebraic

$$\text{Proj}_{\vec{v}_i}(\vec{x}+\vec{y}) = \frac{v_i \cdot (x+y)}{v_i \cdot v_i} v_i = \frac{v_i \cdot x}{v_i \cdot v_i} v_i + \frac{v_i \cdot y}{v_i \cdot v_i} v_i$$

$$= \text{Proj}_{\vec{v}_i}(\vec{x}) + \text{Proj}_{\vec{v}_i}(\vec{y})$$

Therefore projection respects vector addition

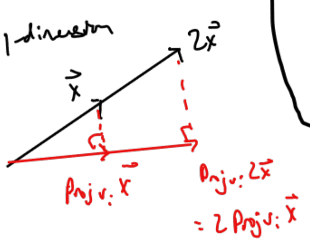
Multiple Dimension

$$\text{Proj}_V(\vec{x}+\vec{y}) = \text{Proj}_{v_1}(\vec{x}+\vec{y}) + \dots + \text{Proj}_{v_k}(x+y)$$

$$= \text{Proj}_{v_1}(x) + \text{Proj}_{v_1}(y) + \dots + \text{Proj}_{v_k}(x) + \text{Proj}_{v_k}(y)$$

$$= \text{Proj}_V(x) + \text{Proj}_V(y)$$

In 1-dimension



Therefore, projection respects scalar multiplication.

$$\text{Proj}_V(c\vec{x}) = c\,\text{Proj}_V(\vec{x})$$

$$\text{Proj}_{v_i} c\vec{x} = \frac{v_i \cdot cx}{v_i \cdot v_i} v_i = c\left(\frac{v_i \cdot x}{v_i \cdot v_i} v_i\right) = c\,\text{Proj}_{v_i}\vec{x}$$

Definition: Let A be and $m \times n$ and B an $n \times p$ matrix defined by

$T_A \circ T_B(\vec{y})$
$= T_A(B\vec{y})$
$= A(B\vec{y})$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np}. \end{bmatrix}$$

Let $T_A : \mathbb{R}^n \to \mathbb{R}^m$ and $T_B : \mathbb{R}^p \to \mathbb{R}^n$ be the linear transformations defined by $T_A(\mathbf{x}) = A\mathbf{x}$ and $T_B(\mathbf{y}) = B\mathbf{y}$. Then the composition $T_A \circ T_B : \mathbb{R}^p \to \mathbb{R}^m$ is a linear transformation, and the associated matrix is AB which we call the **matrix product** of A and B.

Theorem 14.3.2: The entries of AB are the dot products of *rows* of A with *columns* of B. If we write

a_1 is vec, treated as row

b_1 is vec, treated as col

$$A = \begin{bmatrix} \text{—} \ \mathbf{a}_1 \ \text{—} \\ \text{—} \ \mathbf{a}_2 \ \text{—} \\ \vdots \\ \text{—} \ \mathbf{a}_m \ \text{—} \end{bmatrix}, \qquad B = \begin{bmatrix} | & | & & | \\ \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_p \\ | & | & & | \end{bmatrix},$$

then we have

$$\boxed{AB = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 & \cdots & \mathbf{a}_1 \cdot \mathbf{b}_p \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 & \cdots & \mathbf{a}_2 \cdot \mathbf{b}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_m \cdot \mathbf{b}_1 & \mathbf{a}_m \cdot \mathbf{b}_2 & \cdots & \mathbf{a}_m \cdot \mathbf{b}_p \end{bmatrix} = \begin{bmatrix} | & | & & | \\ A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p \\ | & | & & | \end{bmatrix}.}$$

More explicitly, the ij-entry of AB is

$$\sum_{k=1}^{n} a_{ik} b_{kj},$$

ith row of A

the dot product of $\mathbf{a}_i$ and $\mathbf{b}_j$. ← jth col of B

Example 2: Let $A = \begin{bmatrix} 2 & 5 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$. Compute AB. Can you compute BA?

$$AB = \begin{bmatrix} 2 & 5 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & -3 & 5 \\ 3 & 0 & 1 \end{bmatrix}$$

2×2 $\quad 2 \times 3 \quad$ 2×3

these need to match to be able to compute

output is row of A and col of B

You can't compute BA bc it's a 2x3 times 2x2. Matrix mult is not commutative.

diagonal matrix bc only diag is nonzero

Let $C = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$. Compute $C^2 = CC$. What is C^n for any positive integer n?

$$C^2 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2^2 & 0 \\ 0 & 1^2 \end{bmatrix}$$

$$C^n = \begin{bmatrix} 2^n & 0 \\ 0 & 1^n \end{bmatrix}$$

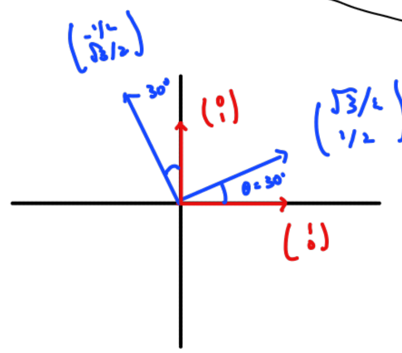
$$C^3 = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2^3 & 0 \\ 0 & 1^3 \end{bmatrix}$$

Example 3: Consider the matrix $A = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$. What does A do geometrically?

Test it on standard basis:

$$A\begin{pmatrix}1\\0\end{pmatrix} = \begin{bmatrix}\sqrt{3}/2\\ 1/2\end{bmatrix}$$

$$A\begin{pmatrix}0\\1\end{pmatrix} = \begin{bmatrix}-1/2\\ \sqrt{3}/2\end{bmatrix}$$



A rotates $\mathbb{R}^2$ by $30°$ counter clockwise

$$A = \begin{bmatrix} \cos(30°) & -\sin(30°) \\ \sin(30°) & \cos(30°) \end{bmatrix}$$

In general, the matrix of counterclockwise rotation of $\mathbb{R}^2$ around the origin by θ is

$$A_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Note that rotations about the origin in $\mathbb{R}^2$ commute: $A_\theta A_\phi = A_\phi A_\theta$.

In $\mathbb{R}^3$, we will focus on rotations around the $x-$, $y-$, and $z-$axes. We have:

- Rotation about the z-axis counterclockwise by angle θ:

keeps z constant, rotate x & y

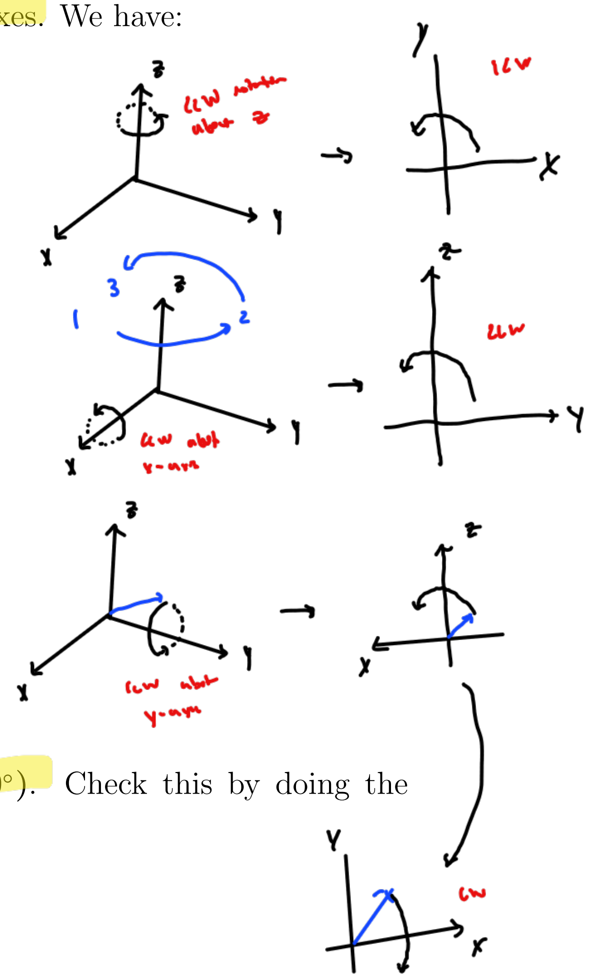
$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cdots \\ \cdots \\ z \end{pmatrix}$$

- Rotation about the x-axis counterclockwise by angle θ:

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} : \begin{pmatrix} x \\ \cdots \\ \cdots \end{pmatrix}$$

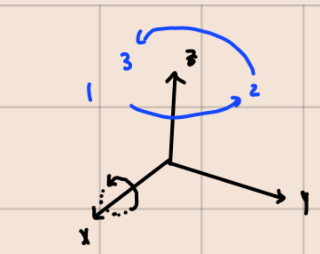
- Rotation about the y-axis counterclockwise by angle θ:

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

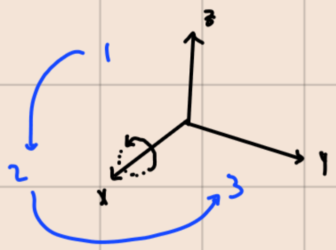


Important: Rotations in $\mathbb{R}^3$, in general, **do not commute**.

Algebraically, for example, $R_x(90°)R_y(90°) \neq R_y(90°)R_x(90°)$. Check this by doing the actual matrix multiplication.



z 90 degrees then x 90 degrees



x 90 degrees then z 90 degrees