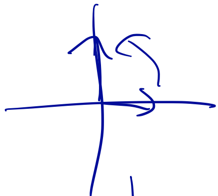
§ 1.9 Matrices of linear transformations

16 Feb

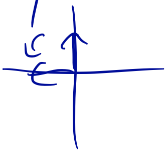
Recall that an $m \times n$ matrix A is determined by where it sends $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$, and these vectors $A\vec{e}_1, \dots, A\vec{e}_n$ are simply the columns of A. This gives a powerful tool for finding the matrix of a given linear transformation: figure out where $\vec{e}_1, \dots, \vec{e}_n$ go and arrange the resulting column vectors into a matrix!

Example Find the matrix of the rotation $90°$ ccw about the origin.

Solution $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ maps to $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$:



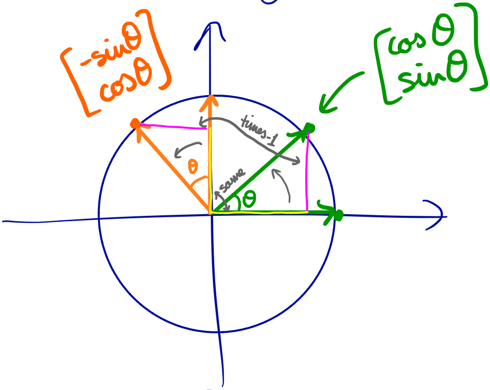
$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ maps to $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$:



So the matrix is $\boxed{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}}$. Done!

Example Find the angle-θ rotation matrix.

Solution



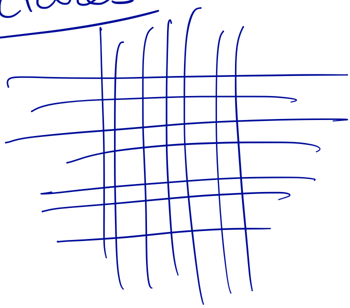
don't worry too much if your trig is shaky. This can be taken as the def^n of sine & cosine!

So we get $\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$.

Note: this is handy for computer graphics, because we can rotate a bunch of points $(x_1, y_1) \ldots, (x_n, y_n)$ at a time calculating and storing the desired rotation matrix and using it to multiply each $\begin{bmatrix}x_i\\ y_i\end{bmatrix}$ for $i = 1, \ldots, n$.

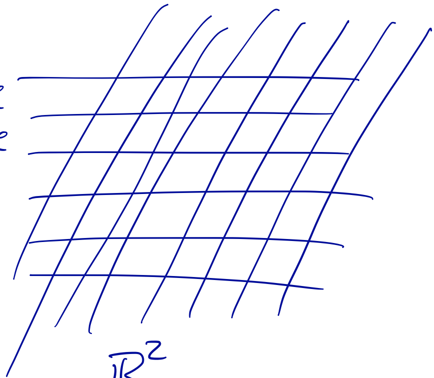
Definitions We say $T: \mathbb{R}^n \to \mathbb{R}^m$ is **surjective** or onto if $\text{range}(T) = \mathbb{R}^m$, i.e. the columns of T span $\mathbb{R}^m$.
And T is **injective**, or one-to-one, if $T(\vec{x}) = T(\vec{y}) \Rightarrow \vec{x} = \vec{y}$.

Pictures

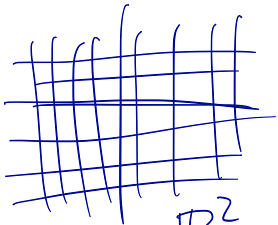


$\mathbb{R}^2$

surjective & injective $\longrightarrow$



$\mathbb{R}^2$



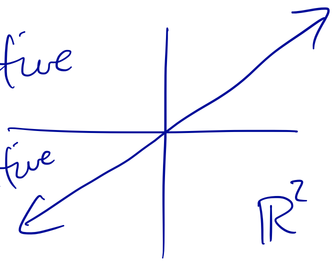
$\mathbb{R}^2$

surjective $\longrightarrow$ not injective

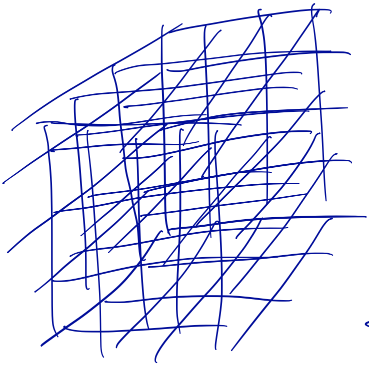
$\mathbb{R}$

$\mathbb{R}^1$

injective $\longrightarrow$ not surjective



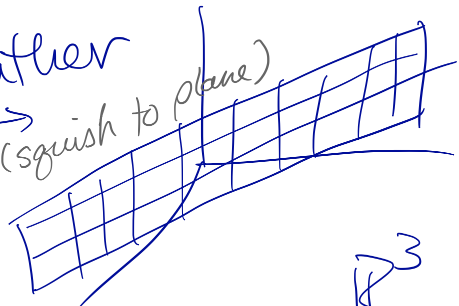
$\mathbb{R}^2$



$\mathbb{R}^3$

neither $\longrightarrow$

(squish to plane)



$\mathbb{R}^3$

Theorem A linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ is injective if and only if $T(\vec{x}) = \vec{0}$ has a unique solution.

The point: we don't need to check uniqueness of solutions for $T(\vec{x}) = \vec{b}$ for all $\vec{b}$, just $\vec{0}$.

Why? if $T(\vec{x}) = 0 = T(\vec{y})$ for $\vec{x} \neq \vec{y}$, then already that's two distinct inputs giving the same output, so T is not injective.

Conversely, if $T(\vec{x}) = T(\vec{y})$ for some $\vec{x} \neq \vec{y}$, then

$$T(\vec{x}) - T(\vec{y}) = 0 \overset{\text{linearity!!}}{\Longrightarrow} T(\vec{x} - \vec{y}) = \vec{0}$$ and so $\vec{x} - \vec{y}$ is a nontrivial "answer" to $T(?) = \vec{0}$. [Recall that a statement "A implies B" is logically equivalent to its contrapositive "not B implies not A". So showing "noninjectivity implies nontrivial solutions" tells us that "no nontrivial solutions implies injectivity".]

Exercise Show that $T(x, y) = (x+y,\ x-y,\ 2x)$ is injective. Solution (idea): Show $\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 2 & 0 \end{pmatrix}$ has a pivot in every column.

!! **Exercise** If $T: \mathbb{R}^n \to \mathbb{R}^m$ is injective and surjective, then $n = m$. (Idea: count pivots)

§2.1 Matrix operations

We define matrix addition and scalar multiplication like for vectors:

have to be the same size

addition
$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 0 & 1 & -1 \\ -2 & 4 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 9 & 9 \end{pmatrix}$$

scalar multiplication
$$3\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 9 \\ 12 & 15 & 18 \end{pmatrix}$$

These operations do what you think:

$$A+B = B+A \quad ; \quad r(sA) = (rs)A$$

$$A+(B+C) = (A+B)+C \quad ; \quad r(A+B) = rA+rB.$$

You can prove any of these if you really want by just writing it out with symbols like $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$, and using corresponding properties of real numbers.

What's more interesting is multiplying two matrices.

Consider that if we want to apply a matrix B to $\vec{x}$ and then multiply the result by A, we'd naturally write

$$\underbrace{A\,\overbrace{B\vec{x}}^{\text{① result of multiplying } B \text{ and } \vec{x}}}_{\text{② then multiply that by } A}.$$

It would be cool if we defined AB so that this expression can be calculated either way: $A(Bx) \overset{\text{we want}}{=} (AB)x$. In function language, we want the matrix product to correspond to the <u>composition</u> of the corresponding linear transformations.

Key idea

So consider:

$$A(B\vec{x}) \overset{\text{def}^{n}\text{ of matrix-vector product}}{=} A\left(x_1\vec{b}_1 + \cdots + x_n\vec{b}_n\right)$$

$$\overset{\text{linearity}}{=} x_1 A\vec{b}_1 + \cdots + x_n A\vec{b}_n$$

$$\overset{\text{def}^{n}\text{ of matrix-vector product, in reverse}}{=} \begin{bmatrix} A\vec{b}_1 & \cdots & A\vec{b}_n \end{bmatrix} \vec{x}$$

The upshot: matrix-matrix product works just like matrix-vector product, one column of B at a time:

Example Find $\underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}}_{A} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$.

Solution:

$$= \begin{bmatrix} A\begin{bmatrix} 1 \\ 5 \\ 9 \end{bmatrix} & A\begin{bmatrix} 2 \\ 6 \\ 10 \end{bmatrix} & A\begin{bmatrix} 3 \\ 7 \\ 11 \end{bmatrix} & A\begin{bmatrix} 4 \\ 8 \\ 12 \end{bmatrix} \end{bmatrix}$$

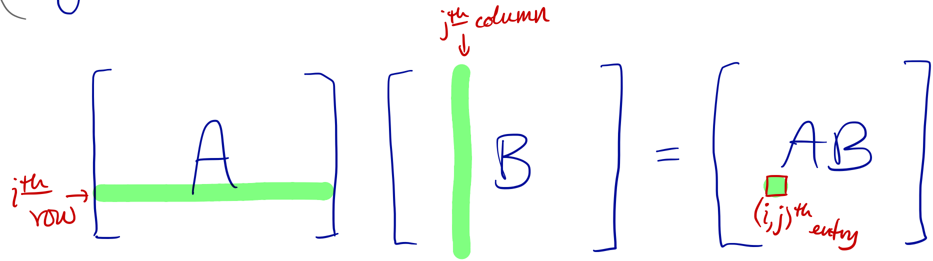
$$= \begin{bmatrix} 38 & 44 & 50 & 56 \\ 83 & 98 & 113 & 128 \end{bmatrix}$$

We can simplify this somewhat:

memorize this. like the actual words. It's worth it

"To find the $(i,j)^{th}$ entry of AB, dot the $i^{\underline{th}}$ row of A with the $j^{\underline{th}}$ column of B"

multiply entries in pairs & sum



Note: If $B: R^p \to R^n$ and $A: R^n \to R^m$, then the composition goes from $\mathbb{R}^p$ to $\mathbb{R}^m$; likewise:

← these have to be equal for composition to make sense.

$$\underbrace{(m \times n)}_{\mathbb{R}^n \to \mathbb{R}^m} \cdot \underbrace{(n \times p)}_{\mathbb{R}^p \to \mathbb{R}^n} = \underbrace{m \times p}_{\mathbb{R}^p \to \mathbb{R}^m}$$

$$\mathbb{R}^n \to \mathbb{R}^m \circ \mathbb{R}^p \to \mathbb{R}^n$$

We define I_n, "the identity matrix", to be the $n \times n$ matrix with 1's down the main diagonal & zeros elsewhere:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In Julia, Python, MATLAB, etc., you do `eye(n)`. lol.

the matrix-matrix product also 'does what you think'

$$(AB)C = A(BC), \quad (rA)B = r(AB), \text{ etc. BUT!!!}$$

Exercise Show that matrix multiplication is <u>not</u> commutative.

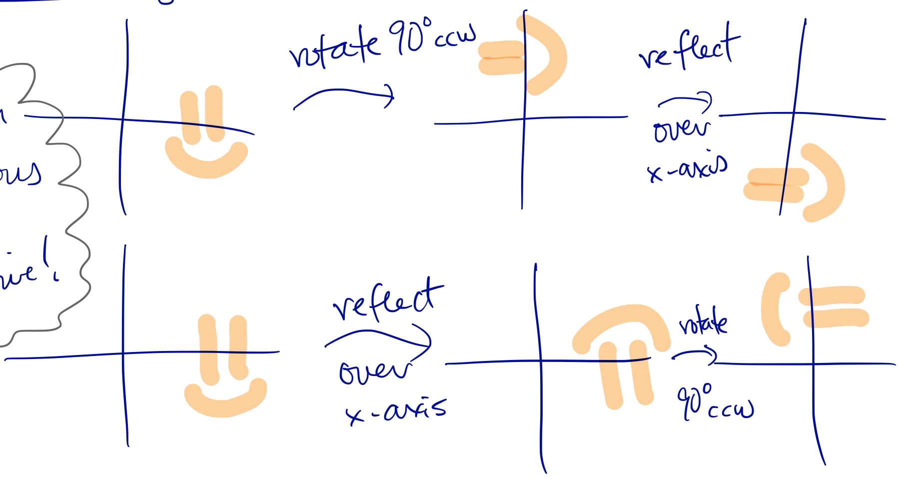
<u>Solution 1</u> Let's choose some numbers and go at it:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix}$$

<u>Solution 2</u> (geometric)

Composition of functions is not commutative!



Definition $A^n = \underbrace{AAA\cdots A}_{n \text{ times}}$. (Only works if A is square)

Definition The transpose of A, denoted A^T, is obtained by forming a matrix whose rows are the columns of A, in the same order.

Example

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

$$2\times 3 \longrightarrow 3\times 2$$

Facts:

$(A+B)^T = A^T + B^T$ (easy)

$(A^T)^T = A$ (easy)

$\longrightarrow$!! $(AB)^T = B^T A^T$ (you can figure it out)

Transpose distributes & reverses order.